

TESE DE DOUTORADO Nº 301

**GALERKIN DESCONTÍNUO NO DOMÍNIO DO TEMPO APLICADO A
PROBLEMAS COM MÚLTIPLAS ESCALAS EM NANOFOTÔNICA**

Fidel Edson de Souza

DATA DA DEFESA: 10/07/2019



Universidade Federal de Minas Gerais
Departamento de Engenharia Elétrica

Programa de Pós-Graduação em Engenharia Elétrica

**Galerkin Descontínuo no Domínio do
Tempo Aplicado a Problemas com
Múltiplas Escalas em Nanofotônica**

Fidel Edson de Souza

Tese de Doutorado

Belo Horizonte
28/06/2019

Universidade Federal de Minas Gerais
Escola de Engenharia
Programa de Pós-Graduação em Engenharia Elétrica

**GALERKIN DESCONTÍNUO NO DOMÍNIO DO TEMPO
APLICADO A PROBLEMAS COM MÚLTIPLAS ESCALAS EM
NANOFOTÔNICA**

Fidel Edson de Souza

Tese de Doutorado submetida à Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do Título de Doutor em Engenharia Elétrica.

Orientador: Prof. Elson José Da Silva

Co-orientador: Prof. Renato Cardoso Mesquita

Belo Horizonte – MG

Julho de 2019

S729g

Souza, Fidel Edson de.

Galerkin descontínuo no domínio do tempo aplicado a problemas com múltiplas escalas em nanofotônica [recurso eletrônico] /Fidel Edson de Souza. - 2019.

1 recurso online (92 f. : il., color.) : pdf.

Orientador: Elson José da Silva.

Coorientador: Renato Cardoso Mesquita.

Tese (doutorado) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Bibliografia: f. 86-92.

Exigências do sistema: Adobe Acrobat Reader.

1. Engenharia Elétrica - Teses. 2. Cristais fotônicos - Teses.
3. Guias de ondas - Teses. 4. Galerkin, Métodos de. - Teses. I. Silva, Elson José da. II. Mesquita, Renato Cardoso. III. Universidade Federal de Minas Gerais. Escola de Engenharia. IV. Título.

CDU: 621.3(043)

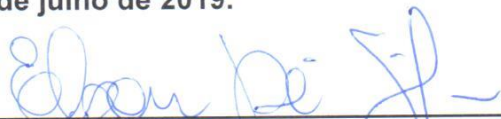
**"Galerkin Descontínuo No Domínio do Tempo Aplicado A
Problemas Com Múltiplas Escalas Em Nanofotônica"**

Fidel Edson de Souza

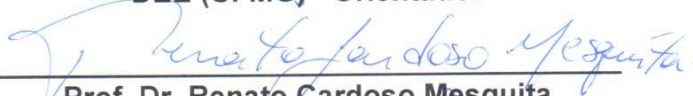
Tese de Doutorado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do grau de Doutor em Engenharia Elétrica.

Aprovada em 10 de julho de 2019.

Por:



Prof. Dr. Elson José da Silva
DEE (UFMG) - Orientador



Prof. Dr. Renato Cardoso Mesquita
DEE (UFMG)



Prof. Dr. Eduardo Henrique da Rocha Coppoli
Engenharia Elétrica (CEFET-MG)



Prof. Dr. Cláudio Garcia Batista
Eng. Telecomunicações (UFSJ -Campus Alto Paraopeba)



Dr. Adriano Chaves Lisboa (Diretor de Pesquisa)
Pesquisa e Desenvolvimento (Gaia Soluções em Engenharia)



Prof. Dr. Cássio Gonçalves do Rego
DELT (UFMG)



Prof. Dr. Ricardo Luiz da Silva Adriano
DEE (UFMG)

Agradecimentos

Primeiramente a Deus pela saúde mental e física, pois sem elas nunca conseguiria realizar esse trabalho.

À minha esposa Elém Patrícia pelo apoio incondicional e compreensão em todos os momentos da nossa vida juntos.

À minha filha Mariana, simplesmente por existir, trazendo muita alegria e completude à minha vida.

Aos meus pais, Marina e Carlúcio, pelo apoio que me deram ao longo da minha caminhada.

À minha família, irmãos sobrinhos e tios, por fazerem parte da minha vida.

À minha sogra Terezinha pela compreensão nos momentos de ausência, e pelas palavras de incentivo.

Aos colegas do LOPAC, Welyson, Roberto, Fernando, Leonardo, Guilherme, Daniele, João, Igor e demais pelos momentos agradáveis no tradicional cafezinho da tarde. Também agradeço pela ajuda que sempre me deram.

Aos professores Elson José da Silva e Renato Cardoso Mesquita pela orientação e por me guiar nos caminhos do conhecimento. Agradeço por terem tido paciência nos momentos difíceis.

"Cem vezes todos os dias lembro a mim mesmo que minha vida interior e exterior, depende dos trabalhos de outros homens, vivos ou mortos, e que devo esforçar-me a fim de devolver na mesma medida que recebi"

—ALBERT EINSTEIN

Resumo

O constante aumento da necessidade de recursos tecnológicos que tragam maiores taxas de transmissão, processamento e armazenamento de dados, impulsiona o estudo das propriedades da propagação eletromagnética em dispositivos de escala nanométrica. Assim, na literatura, são encontrados uma grande quantidade de trabalhos científicos com essa finalidade. Com os avançados *hardwares* que disponibilizamos nos dias atuais, juntamente com o desenvolvimento crescente dos métodos numéricos, cada vez mais os estudos através de modelos numéricos ganham força entre os pesquisadores em todo o mundo.

Um dos dispositivos fotônicos mais promissores é o guia planar, baseado em cristais fotônicos bidimensionais. Métodos numéricos no domínio do tempo podem simular a propagação nestes guias revelando características peculiares, como a luz lenta. Dentre os problemas encontrados na modelagem de dispositivos nanofotônicos, estão os problemas com múltiplas escalas. Os métodos numéricos têm dificuldade de se ajustarem as diferentes escalas. O Galerkin descontínuo no domínio do tempo (*discontinuous Galerkin time domain*, DGTD) é um método promissor no tratamento de problemas com múltiplas escalas, pois usa malhas não estruturadas na discretização do domínio. Porém, na versão convencional, a integração do tempo pode trazer um grande custo computacional. Por isso, as estratégias de passo de tempo local (LTS - *local time stepping*) têm sido propostas por alguns pesquisadores. Apesar dos métodos LTS existentes serem eficientes, ainda apresentam limitações e não exploram todo o potencial da discretização espacial DG. Portanto, ainda há espaço para o desenvolvimento de métodos ainda mais eficientes. Dessa forma, nosso objetivo é estudar e desenvolver um método LTS-DGTD eficiente no tratamento de problemas de nano fotônica com múltiplas escalas.

A estratégia LTS desenvolvida nesse trabalho é baseada no método linear com multipassos com forte preservação da estabilidade (SSPMS). Contudo, a princípio, pode ser aplicada à qualquer método que possua um único estágio. Os testes foram realizados em problemas de propagação eletromagnética, inclusive em guias planares baseados em cristais fotônicos. Os resultados obtidos validam a estratégia proposta e demonstram sua eficiência.

Palavras-chave: cristais fotônicos; guias de ondas; multiescala; passo de tempo local; LTS-DGTD.

Abstract

The constant increase in the need for technological resources that bring higher rates of transmission, processing, and storage of data, drives the study of the properties of electromagnetic propagation in nanoscale devices. Thus, we find in the literature a significative number of scientific researches for this purpose. With the advanced hardware available nowadays and the increasing development of numerical methods, research using numerical models simulations are becoming more common around the world.

One of the most promising photonic devices is the planar guide, based on two-dimensional photonic crystals. Numerical methods in the time domain can simulate propagation in these guides revealing peculiar characteristics, such as slow light. Among the problems encountered in the modeling of nanophotonic devices are problems with multiple scales. Numerical methods have difficulty adjusting to different scales. The DGTD is a promising method in the treatment of problems with multiple scales since it uses unstructured meshes in domain discretization. However, in the standard version, the time integration can bring a big computational cost. Therefore, researchers have been proposed local time stepping strategies (LTS). Although the existing LTS methods are efficient, they still have limitations and do not exploit the full potential of spatial discretization DG. Therefore, it is possible to develop more efficient LTS strategies.

The LTS strategy developed here is based on the linear multistep strong stability preserving method (SSPMS). However, in principle, it can be applied to any single stage method. To test the strategy, we applied on electromagnetic wave propagation on photonic crystals planar guides. The results validate and demonstrate the multiclass strategy efficiency.

Keywords: photonic crystals; wave guides; multiscale; local time stepping; LTS-DGTD.

Sumário

Lista de Símbolos e Abreviaturas	viii
Lista de Figuras	xii
Lista de Tabelas	xiii
1 Introdução	1
1.1 Apresentação	1
1.1.1 Cristais Fotônicos	2
1.1.2 Guias Planares	2
1.2 Estado da Arte	6
1.2.1 Modelagem no Domínio da Frequência	6
1.2.1.1 Estruturas de Bandas	6
1.2.2 Simulação no Tempo de Problemas com Multiescala em Nanofotônica	7
1.2.2.1 LTS-DGTD	10
1.3 Contribuições da Tese	13
2 Fundamentação Teórica	14
2.1 Fundamentos de Cristais Fotônicos	14
2.1.1 Equações de Maxwell em Meios Dielétricos	14
2.1.2 Simetrias relacionadas ao Cristal Fotônico	16
2.1.2.1 Inversão	16
2.1.2.2 Simetria Translacional Contínua	16
2.1.2.3 Simetria Translacional Discreta	17
2.1.2.4 Estrutura de Bandas Fotônicas	19

2.1.2.5	Simetria Rotacional e Zona Irredutível de Brillouin	20
2.1.2.6	Simetria Especular e Separação de Modos	20
2.1.3	Cristal Fotônico Bidimensional	20
2.1.3.1	Modos de Bloch Bidimensionais	21
2.1.3.2	Rede Quadrada de Colunas Dielétricas	22
2.1.3.3	Defeitos Lineares e Guias de Ondas	23
2.1.3.4	Projetando Guias de Cristais 2D	25
2.2	Modelagem das Bandas de Cristais Fotônicos 2D pelo Método de Galerkin	26
2.2.1	Forma Forte	26
2.2.1.1	Equação governante	26
2.2.1.2	Condições Quasi-Periódicas	27
2.2.2	Forma Fraca	28
2.2.2.1	Modo TE	28
2.2.2.2	Modo TM	29
2.2.3	Discretização	29
2.2.4	Funções de Base	30
2.2.4.1	Funções NEM	30
2.2.4.2	Diagrama de Voronoi	30
2.2.4.3	Triangulação Dual de Delaunay	31
2.2.4.4	Célula de Voronoi de Segunda Ordem	31
2.2.4.5	Construção das Funções de Forma	32
2.2.5	Imposição das Condições Periódicas	32
2.2.6	Validação Computacional	35
2.3	Simulações no Domínio do Tempo	37
3	Galerkin Descontínuo no Domínio do Tempo	39
3.1	Equações de Maxwell	39
3.2	Lei de Conservação	40

3.2.1	Modo TM	41
3.2.2	Modo TEM	42
3.3	Forma Variacional	42
3.3.1	Fluxo Numérico	43
3.3.1.1	Condições de Fronteira	44
3.3.1.2	Fontes	45
3.4	Forma Semi-discreta	46
3.4.1	Modo TM	48
3.4.2	Modo TEM	52
3.4.2.1	Testes e Resultados	54
3.5	Integração do Tempo	55
3.5.1	Método <i>leapfrog</i> - LF	56
3.5.2	Método Runge-Kutta	57
3.5.3	Estabilidade Condicional	58
3.5.4	Passo de Tempo Local - LTS	58
3.5.4.1	<i>Causal-Path Local Time Stepping</i> - CPLTS	58
3.5.4.2	LF2-CPLTS	59
3.5.4.3	LSRK4-CPTLS	59
3.6	Método multi passo linear com forte estabilidade - SSPMS	61
3.6.1	Método multi classe <i>mC</i> -SSPMS	63
4	Testes Numéricos	66
4.1	Equação de Advecção	66
4.2	Propagação Eletromagnética	69
4.2.1	Teste 1: Propagação de um pulso gaussiano	69
4.2.2	Teste 2: Linha de corrente	75
4.2.3	Teste 3: Guia de onda em L baseado em cristal fotônico	76
4.2.4	Teste 4: Porta lógica <i>or</i>	79

5 Conclusão e Trabalhos Futuros	84
5.1 Conclusão	84
5.2 Trabalhos Futuros	85
Referências Bibliográficas	86

Lista de Símbolos e Abreviaturas

$+$	referência ao elemento vizinho
$-$	referência ao elemento local
Δt	passo de tempo
Δx	menor divisão de um <i>grid</i>
$\hat{\mathbf{n}}$	vetor unitário normal à fronteira
λ	comprimento de onda
$\mathcal{C}$	coeficiente SSP
$\mathcal{D}_r$	operador diferencial em relação à variável r
$\mathcal{D}_s$	operador diferencial em relação à variável s
$\mathcal{K}$	matriz de rigidez
$\mathcal{M}$	matriz de massa
$\mathcal{V}$	matriz de Vandermonde generalizada
$\mathcal{V}_{r,(i,j)}$	derivada da matriz de Vandermonde em relação à r
$\mathcal{V}_{s,(i,j)}$	derivada da matriz de Vandermonde em relação à s
$\mathcal{C}_r$	fator de redução
μ	permeabilidade magnética
μ_0	permeabilidade no vácuo
ω	frequência angular
ϕ_m	parâmetro de uniformidade entre classes
$\mathbf{B}$	campo de indução magnética
$\mathbf{b}$	vetor recíproco de rede

D	campo de deslocamento elétrico
E	vetor campo elétrico
F*	fluxo numérico
H	vetor campo magnético
k	vetor de onda
q_k	solução para o elemento local <i>k</i>
R	vetor de rede do cristal fotônico
r	vetor posição
Θ	operador diferencial associado ao cristal fotônico
ε	permissividade elétrica
ε₀	permissividade elétrica no vácuo
ε_r	constante dielétrica em função de r
a	constante de rede de um cristal fotônico
cm	centímetro
dB	decibéis
e_r	erro relativo
E_z	componente <i>z</i> do campo elétrico
f_c	fator de concentração de campo elétrico
G_p	ganho percentual
H¹	espaço de Hilbert, onde as funções possuem derivadas primeiras com quadrado integrável
h^e	comprimento da aresta do triângulo
H_z	componente <i>z</i> do campo magnético
hp – FEM	FEM com adaptatividade em termos das dimensões do elemento (<i>h</i>) e ordem polinomial (<i>p</i>)
k_x	componente <i>x</i> do vetor de onda

k_y	componente y do vetor de onda
k_z	componente z do vetor de onda
mC-SSPMS	multiclass linear multistep strong stability preserving method
N_{el}	número de elementos
$p_1 - FEM$	FEM com ordem polinomial igual a 1
P_m	polinômios de Legendre
$T_{\mathbf{R}}$	operador translacional
t_{exec}	tempo de execução
Y	admitância do meio
Z	impedância do meio
c	velocidade de uma onda
CFL	condição de Courant–Friedrichs–Lewy
CPLTS	<i>causal path local time stepping</i>
CPU	<i>central processing unit</i>
DFT	<i>Discrete Fourier Transform</i>
DG	<i>discontinuous Galerkin</i>
DG-FEM	<i>discontinuous galerkin finite element method</i>
DGTD	<i>discontinuous galerkin time domain</i>
FDTD	<i>finite difference time domain</i>
FEM	<i>finite element method</i>
FETD	<i>finite element time domain</i>
FVM	<i>finite volume method</i>
G-FEM	<i>generalized finite element method</i>
GFEM	<i>generalized finite element method</i>
GVD	<i>Group Velocity Dispersion</i>

IMEX	<i>implicit-explicit</i>
LF	<i>leapfrog</i>
LF-LTS	<i>leapfrog</i> com passo local de tempo
LF2	<i>leapfrog</i> de segunda ordem
LSERK4	low storage explicit four order runge kutta
LTS	<i>local time stepping</i>
NEM	<i>natural element method</i>
PEC	<i>Perfect Electrical Conductor</i>
PhC	<i>photonic crystal</i>
PMC	<i>Perfect Magnetic Conductor</i>
PML	<i>Perfectly Matched Layer</i>
PWE	<i>plane wave expansion method</i>
RK4	runge kutta de quarta ordem
S-FEM	<i>Spectral Finite Element Method</i>
TE	transverso elétrico
TEM	transverso elétrico e magnético
TM	transverso magnético
TVD	<i>total variation diminishing</i>
X-FEM	<i>extended finite element method</i>

Lista de Figuras

1.1	Guia de ondas baseado em um cristal planar formado por furos de ar em uma membrana de silício (Baba (2008))	3
1.2	Representação de um guia de ondas baseado em um cristal fotônico bidimensional formado por pilares dielétricos imersos em ar	4
1.3	Malha não estruturada formada por elementos triangulares, representando o guia de ondas da figura 1.2	4
2.1	Simetria Translacional Discreta	17
2.2	Cristais Fotônicos Bidimensionais	21
2.3	Estrutura de bandas de um cristal bidimensional de rede quadrada de colunas dielétricas com constante $\epsilon_r = 8,9$ e raio $r/a = 0,2$.	23
2.4	Guia W1 formado por uma linha de defeitos em um cristal fotônico bidimensional com rede quadrada de colunas dielétricas imersas no ar.	24
2.5	Estrutura de bandas projetada e diagrama de dispersão de um modo guiado em uma linha de defeitos de largura d .	24
2.6	Modo TM propagando em um guia W1, formado por uma linha de defeitos com largura $d = 0,912\mu m$.	26
2.7	Representação topológica das fronteiras da célula unitária.	27
2.8	(a): Diagrama de Voronoi de primeira ordem de um conjunto de M nós. (b): Triangulação dual de Delaunay e circuncírculos. (c): Célula de Voronoi de segunda ordem de x .	31
2.9	Representação topológica da distribuição dos nós no domínio com duas células adjacentes.	33
2.10	Estruturas de Bandas. (a): Estrutura de bandas de um cristal formado de colunas dielétricas com constante $\epsilon_r = 8,9$ e raio $r = 0,2a$. (b): Estrutura de bandas de um cristal formado de veias dielétricas com constante $\epsilon_r = 8,9$ e espessura $l = 0,165a$.	36

2.11	Energia: Distribuição da energia de campo elétrico $\epsilon_r \mathbf{E}^2(x,y)$ de um domínio formado por 9 células unitárias.	37
3.1	Propagação eletromagnética de um modo TEM_x no ar. O numero de elementos usados foi $N_e = 401$ e o grau do polinômio de Lagrange $p = 6$.	55
3.2	Comparação entre os fluxos central e <i>upwind</i> . O numero de elementos usados foi $N_e = 2001$.	55
3.3	Passo local no LF2-CPTLS com $h_1 = 1/2$ e $h_2 = 1/2$. Angulo <i>et al.</i> (2014)	60
3.4	Evolução de classes.	64
4.1	Equação de Advecção: testes experimentais.	68
4.2	Equação de Advecção: erro relativo.	69
4.3	Malha multiescala triangular para resolver o problema de propagação de pulso.	70
4.4	Teste 1: propagação do pulso gaussiano com quatro soluções em instantes diferentes.	70
4.5	Parâmetros de não uniformidade entre classes.	73
4.6	Malha teste 2 com 4896 elementos.	75
4.7	Teste 2: solução para o campo elétrico.	76
4.8	Erro relativo <i>versus</i> ordem polinomial para o teste 2.	76
4.9	(a) Malha multiescala representando um guia de ondas de cristal fotônico. (b) Solução E_z para o 3C-SSPMS(6,3) com $N = 5$. (c) Vetor de Poynting.	78
4.10	Dispersão sofrida pelo pulso ao se propagar pelo guia de ondas	79
4.11	(a): representação geométrica da porta <i>or</i> e da região de PML que a cerca. (b): malha não estruturada representando o domínio computacional.	80
4.12	Campo elétrico, E_z , para teste 3 usando o esquema 3C-SSPMS(6,3) com $N = 4$. A janela de tempo utilizada foi de $50\mu s$. (a) solução para $A = 1$, $B = 0$. (b) solução para $A = 0$ e $B = 1$. (c) solução para $A = 1$ e $B = 1$.	81
4.13	Parâmetro de não uniformidade entre classes.	82
4.14	Potência instantânea	83

Lista de Tabelas

3.1	Condições de fronteira	45
4.1	Equação de Advecção: fator de redução C_r	68
4.2	Teste 1: Divisão de elementos entre classes	71
4.3	Teste 1: eficiência e precisão de mC -SSPMS(6,3), SSPMS(6,3) e RK4	71
4.4	Teste 1: Fator C_r para o mC -SSPMS(6,3) com interpolação	72
4.5	Teste 1: eficiência e acurácia LF2-CPLTS	73
4.6	Fator C_r LF2-CPLTS	74
4.7	Teste 1: eficiência e acurácia do mC -SSPMS(6,3) com subclasses	74
4.8	Teste 1: fator de redução C_r para o mC -SSPMS(6,3) com subclasses	74
4.9	Teste 2: eficiência e precisão de mC -SSPMS(6,3), SSPMS(6,3) e RK4	76
4.10	Teste 3: Divisão de elementos entre classes	77
4.11	Teste 3: eficiência e acurácia para mC -SSPMS(6,3), SSPMS(6,3) and RK4.	78
4.12	Teste 4: Resultados da simulação dos esquemas mC -SSPMS(6,3), SSPMS(6,3)	82

CAPÍTULO 1

Introdução

1.1 Apresentação

A luz é uma importante forma de interação entre o meio ambiente e o homem. Há tempos a humanidade procura desenvolver maneiras de controlá-la e usá-la no desenvolvimento de tecnologias em diversas áreas da ciência e da engenharia. A partir do estudo da fotônica foi possível construir dispositivos como a fibra óptica, que proporcionaram um grande aumento das taxas de transmissão de dados em redes de comunicação. Recentemente, muitos pesquisadores têm concentrado seus esforços no estudo da fotônica em escalas de micrômetros e nanômetros. Dessa forma, atualmente já se tem ciência de muitos dispositivos tecnológicos capazes de interagir com a luz em escalas nanométricas. Muitos destes dispositivos são baseados em Cristais Fotônicos (*Photonic Crystals* - PhC).

Existem basicamente duas formas de estudar estas estruturas nanométricas: teórica e experimental. Estudos teóricos são importantes, pois, permitem conhecer aspectos da estrutura a partir de uma modelagem analítica ou numérica. Normalmente, é muito difícil encontrar uma representação analítica para uma estrutura mais realística, e por isso os métodos numéricos são muito usados em estudos teóricos. Os mais usados no estudo dos cristais fotônicos são: expansão em ondas planas (*plane wave expansion method* - PWE) e o método de diferenças finitas no domínio do tempo (*finite difference time domain* - FDTD). Apesar de eficientes, esses métodos apresentam limitações. Por isso, recentemente métodos variacionais, baseados em malhas não estruturadas, têm encontrado mais adeptos entre a comunidade científica. Entre os mais importantes estão: o método de elementos finitos (*finite element method* - FEM) e o Galerkin descontínuo no domínio do tempo (*discontinuous Galerkin time domain* - DGTD).

O método numérico deve ser escolhido de acordo com o problema a ser tratado. Em nanofotônica existem algumas classes de problemas e cada uma tem suas características particulares. Portanto, é importante conhecer as características do problema que se pretende resolver. Basicamente, existem quatro classes de problemas em nanofotônica: propagação eletromagnética, confinamento de modos, espalhamento da luz e multiescala. Em modelos mais realísticos podemos ter combinações de classes. No estudo da propagação, o método mais usado é o FDTD. Porém, encontra limitações no tratamento dos problemas de propagação em um dispositivo com múltiplas escalas. Neste cenário o método DGTD se mostra mais preciso e eficiente

(Chen e Liu (2013)). Guias de ondas com base em cristais fotônicos possuem vários espalhadores com tamanho que se aproximam do comprimento de onda da luz propagante. Como será demonstrado nesse trabalho, o método DGTD possui características que são adequadas para a modelagem deste dispositivo.

1.1.1 Cristais Fotônicos

A transmissão de grandes volumes de informação, por longas distâncias e com baixas perdas, tem revolucionado a indústria da comunicação. Isto é conseguido, nos dias de hoje, principalmente com a utilização das fibras ópticas. Entretanto, há décadas vêm se desenvolvendo estruturas que têm mudado radicalmente a forma com que se controla o fluxo da luz em uma escala microscópica e nanoscópica. Estas estruturas são conhecidas como cristais fotônicos. Dessa forma, desenvolver dispositivos ópticos, cuja capacidade, velocidade e largura de banda apresentem potencial elevado, é um dos grandes objetivos da engenharia elétrica atualmente.

Os cristais fotônicos são estruturas que apresentam periodicidade de seu índice de refração em pelo menos uma dimensão, possibilitando assim a proibição total ou parcial da propagação luminosa em certas direções de interesse. Cristais fotônicos unidimensionais possuem constante dielétrica periódica ao longo de uma direção e homogênea nas demais. Já os cristais fotônicos bidimensionais possuem constante dielétrica periódica ao longo de um plano, e homogênea na terceira dimensão. Os cristais tridimensionais possuem constante dielétrica periódica ao longo das três dimensões.

Assim como nos semicondutores, que possuem bandas proibidas (*band gaps*) à condução de elétrons, nos cristais fotônicos existem bandas proibidas à propagação luminosa. Estas regiões são seletivas em frequência, ou seja, algumas faixas de frequência não podem propagar no interior da estrutura. As faixas do espectro de frequência onde ocorrem as bandas proibidas são determinadas pelas características físicas do cristal. Portanto, é possível criar *gaps* fotônicos específicos para uma faixa do espectro. Esta propriedade permite o uso destas estruturas como base de vários dispositivos tecnológicos. Alguns exemplos são: espelhos de Bragg, guias de onda, cavidades ressonantes, fibras de cristais fotônicos e até transistores ópticos (Joannopoulos *et al.* (2011); Sibilía *et al.* (2008)).

1.1.2 Guias Planares

Os cristais fotônicos são considerados estruturas promissoras capazes de alavancar a tecnologia de circuitos ópticos integrados (Baba (2007)). Atualmente, os guias de onda com base em cristais planares (membranas) figuram entre as tecnologias ópticas integradas mais bem sucedidas. Isto é devido às suas baixas perdas de propagação, na ordem de 10 dB/cm

(Dulkeith *et al.* (2005)). É possível projetar muitos dispositivos passivos que exploram as propriedades ópticas únicas deste tipo de guia. Propriedades tais como: luz lenta ou ajuste da dispersão (Kappeler *et al.* (2011)).

Guias baseados em cristais planares confinam a luz em uma direção usando o *gap* fotônico e na outra por reflexão total. Para isso, é necessário que o guia seja cercado por um material de menor índice de refração (Villeneuve *et al.* (1998)). Estes guias são construídos inserindo uma, ou mais linhas de defeitos em um cristal fotônico planar, que é uma estrutura 3D, finita, formada por um cristal com periodicidade bidimensional. Um exemplo deste guia é mostrado na Figura 1.1, encontrada em Baba (2008). Estes guias operam normalmente em comprimentos de ondas na faixa do infravermelho.

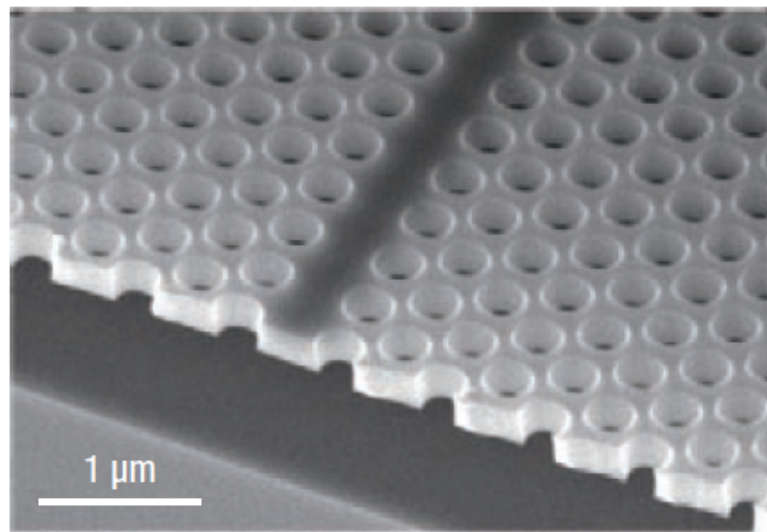


Figura 1.1: Guia de ondas baseado em um cristal planar formado por furos de ar em uma membrana de silício (Baba (2008))

Guias de ondas baseados em cristais fotônicos possuem características dispersivas muito peculiares, que podem ser encaradas como um desafio a transmissão de sinais. Por outro lado, podem dar origem a fenômenos únicos, como a luz lenta. A luz com uma velocidade de grupo notavelmente baixa, é uma solução promissora para o armazenamento (*buffering*) e processamento no domínio do tempo, dos sinais ópticos. Também oferece a possibilidade de compressão espacial de energia óptica e o aumento de efeitos ópticos, lineares e não-lineares (Baba (2008)). Esta característica pode ser usada na idealização e construção de diversos dispositivos. Por exemplo em Soljačić *et al.* (2002), é demonstrado que uma baixa velocidade de grupo pode aumentar dramaticamente as mudanças de fase induzidas, causadas por pequenas alterações no índice de refração. Isto pode ser utilizado para diminuir o tamanho de muitos dispositivos, incluindo: comutadores, roteadores, portas lógicas, conversores de comprimento de onda, dentre outros. As características dispersivas únicas deste tipo de guia podem propiciar a propagação de sólitons, ou seja envelopes com dispersão (GVD-*group velocity dispersion*) nula, como pode ser visto em Colman *et al.* (2010).

O estudo teórico destes guias de onda é de muito interesse, pois pode prover um entendimento sobre os fenômenos associados e ajudar a projetar estruturas que possuam os requisitos desejados, antes de partir para a construção de um protótipo. A fim de estudar dispositivos baseados em guias planares, muitos autores optam por trabalhar com uma forma mais simplificada, os guias bidimensionais (Karle *et al.* (2004), Mori e Baba (2004)). A Figura 1.2 apresenta um cristal fotônico bidimensional formado por pilares dielétricos imersos em ar.

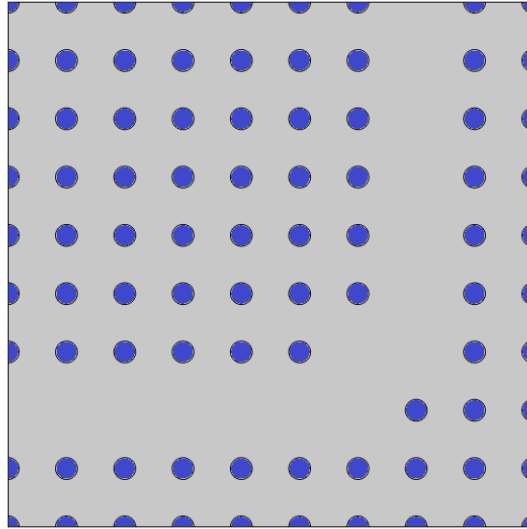


Figura 1.2: Representação de um guia de ondas baseado em um cristal fotônico bidimensional formado por pilares dielétricos imersos em ar

Um importante aspecto do estudo teórico é a determinação dos métodos a serem utilizados. Estes métodos podem ser divididos em duas categorias, domínio da frequência e domínio do tempo. Apesar de distintos, devem apresentar algumas características básicas em comum: alta convergência, precisão e eficiência computacional. Por muitas vezes o cristal fotônico base para o guia pode ser formado por arranjos dielétricos arredondados, como os pilares mostrados na Figura 1.2. Isso dificulta a representação da geometria por meio de grades uniformes, sendo mais adequado representá-la a partir de uma malha não estruturada como na Figura 1.3.

A malha não estruturada permite que algumas regiões sejam mais refinadas que outras. Isso permite representar com mais eficiência o domínio físico. Contudo, com a discrepância das dimensões dos elementos que formam o domínio computacional surge o problema da multiescala. Em simulações no domínio do tempo com métodos explícitos, a multiescala pode trazer esforço computacional desnecessário, se não for tratada adequadamente. Portanto, nesse trabalho serão desenvolvidas e analisadas formas de tratar problemas com multiescala usando o método DGTD.

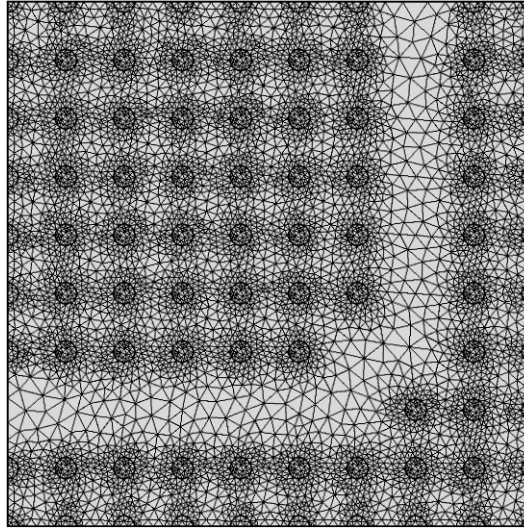


Figura 1.3: Malha não estruturada formada por elementos triangulares, representando o guia de ondas da figura 1.2

1.2 Estado da Arte

O trabalho desenvolvido nessa tese é dividido em duas partes. Primeiramente, uma modelagem no domínio da frequência baseada no método dos elementos naturais, NEM - *natural element method*, para o cálculo do diagrama de estrutura de bandas do cristal base é desenvolvida. Por fim, um método com passo local de tempo, LTS-DGTD, para simular a propagação eletromagnética em dispositivos baseados em cristais fotônicos é proposto e desenvolvido.

1.2.1 Modelagem no Domínio da Frequência

A modelagem matemática usada para o estudo dos cristais fotônicos é obtida a partir das equações de Maxwell, resultando em equações diferenciais parciais, de primeira ou segunda ordem, escalares ou vetoriais, dependendo da abordagem escolhida. Esses modelos geralmente não possuem soluções analíticas, sendo necessário o uso de métodos numéricos para encontrar tais soluções. Entre as diferentes técnicas numéricas disponíveis na literatura para modelagem de cristais fotônicos estão: expansão em ondas planas, diferenças finitas no domínio do tempo, o método dos elementos finitos, além de suas variações (FEM, (Generalizado) GFEM, (Estendido) X-FEM, (Descontínuo) DG-FEM). Atualmente a literatura também apresenta modelagem de cristais fotônicos usando métodos sem malhas e o método dos elementos naturais. Uma revisão de métodos numéricos aplicados à nanofotônica é apresentada em [Gallinet *et al.* \(2015\)](#).

1.2.1.1 Estruturas de Bandas

No estudo de estruturas de bandas de cristais fotônicos, o método mais usado tem sido expansão em ondas planas. Apesar do sucesso inicial desse método, ele possui sérias limitações computacionais. Primeiramente, quando a permissividade elétrica não é uma função analítica, a convergência do método pode ser muito lenta, pois é necessário o uso de DFT's (*Discrete Fourier Transform*). Um outro problema surge quando o método é aplicado a estruturas finas com alto índice de refração, configuração que favorece a formação das bandas proibidas. Devido a reflexão total interna, surgem modos de baixa frequência. Tais modos são muito difíceis de serem determinados usando o método de expansão em ondas planas. Um outro problema encontrado neste método é que as matrizes associadas a ele são inerentemente cheias. Isto torna os algoritmos muito custosos computacionalmente, e muitas vezes proibitivos (Axmman e Kuchment (1999)).

O FDTD também é utilizado na obtenção de estruturas de bandas de cristais fotônicos. Como trabalha no domínio do tempo, possibilita o cálculo de uma faixa larga de frequências de uma só vez. Contudo, é necessário a aplicação de algum método para levar a solução para o domínio da frequência, como Fourier. Isso pode acarretar no aparecimento de modos espúrios (Joannopoulos *et al.* (2011)). Além disso, esse método utiliza malhas estruturadas na discretização do domínio, ou seja, na forma convencional, a malha do FDTD é um *grid* retangular. Isto faz com que ela não se adapte bem a certos domínios ou contornos, o que é um problema, principalmente se tratando de cristais bidimensionais e tridimensionais.

O método dos elementos finitos já vem sendo utilizado no estudo de cristais fotônicos como pode ser visto em Axmann e Kuchment (1999); Burger *et al.* (2005); Contu *et al.* (2012). O uso do FEM se deve a sua eficiência comprovada, e algumas características desejáveis, como a possibilidade de tratar facilmente as descontinuidades encontradas na função permissividade elétrica. As dimensões dos elementos da malha podem ser ajustadas próximo às regiões onde se espera singularidades na solução. As matrizes associadas ao método são esparsas, o que diminui consideravelmente o custo computacional. A malha de elementos finitos é não estruturada, permitindo a ela se ajustar a qualquer contorno ou descontinuidade do domínio com relativa facilidade. Além disso existem versões deste método que podem trazer uma alta convergência, como elementos espectrais (S-FEM - *Spectral Finite Element Method*) de alta ordem e também o *hp* - FEM. Adicionalmente, usando condições periódicas é possível reduzir o domínio computacional a uma célula unitária reduzindo de forma significativa o custo computacional.

Um outro método que se apresenta como opção para a obtenção de diagramas de bandas de cristais bidimensionais é o método dos elementos naturais. O NEM possui algumas boas características do FEM e algumas boas características dos métodos sem malha. Possui uma formulação semelhante ao FEM tendo como principal diferença as funções usadas na discreti-

zação da solução. O NEM é baseado na forma fraca global e no método de Galerkin. Utiliza como funções de forma os *natural neighbors coordinates*, desenvolvidos por [Sibson \(1980\)](#) para interpolação de dados. Sendo essas funções interpolantes, as condições essenciais podem ser inseridas diretamente, de forma forte. Elas também possuem consistência de ordem 0 e 1, ou seja, conseguem reproduzir exatamente polinômios de grau 1 e constantes ([Sukumar \(1998b\)](#)). Apesar das funções do NEM serem construídas com base em uma triangulação, a precisão da aproximação não depende da qualidade dos triângulos, que é uma característica dos métodos sem malha. Além disso, o NEM pode oferecer maior precisão que o p_1 -FEM tradicional, quando aplicado a problemas eletromagnéticos ([Bruyère et al. \(2008\)](#)).

1.2.2 Simulação no Tempo de Problemas com Multiescala em Nanofotônica

Um problema de multiescala é caracterizado quando partes do domínio apresentam dimensões muito distintas. Esse é o caso de um guia de ondas com base em um cristal fotônico bidimensional. Para simular uma propagação por um guia deste tipo, com precisão, é necessário que as geometrias do domínio sejam bem representadas. Quando essa simulação é feita no domínio do tempo, isso pode resultar em um grande custo computacional.

O método numérico mais utilizado na simulação de dispositivos baseados em cristais fotônicos, no domínio do tempo, é o FDTD proposto por [Yee et al. \(1966\)](#). Isso se deve à sua simplicidade e eficiência. Existem muitos exemplos do uso do FDTD no projeto e análise de dispositivos baseados em cristais fotônicos, por exemplo: [Painter et al. \(1999\)](#) e [Sinha et al. \(2011\)](#). Apesar do método FDTD apresentar eficiência, ainda apresenta limitações. Este método discretiza o espaço através de *grids* uniformes estruturados e por isso não se ajusta bem a contornos arredondados, causando erros de *staircase*. O espaçamento do *grid*, Δx , é crítico para uma alta convergência. Em problemas de propagação é recomendado que $\Delta x \leq \lambda/20$ ([Kahnert \(2003\)](#)), sendo λ o comprimento de onda. Em nanofotônica os comprimentos de onda estão normalmente na ordem de micrômetros ou nanômetros. Isso implica em um espaçamento muito pequeno. A integração no tempo no FDTD é feita utilizando o algoritmo *leapfrog* e para garantir a estabilidade é necessário respeitar a condição de Courant ([Taflove et al. \(1995\)](#)), que obriga o passo de tempo, Δt , ser da mesma ordem que $\Delta x/c$, sendo c a velocidade da onda no meio. A uniformidade do *grid* é um fator que prejudica a eficiência no tratamento de problemas com múltiplas escalas, pois é necessário se ajustar às menores, levando a um esforço computacional desnecessário.

Um outro método de simulação no tempo muito usado, é o elementos finitos no domínio do tempo (*finite element time domain* - FETD). Nele é usada uma malha não estruturada. Dessa forma, os contornos e as múltiplas escalas podem ser melhor representados. Porém, a integração do tempo é feita de forma implícita. Isso significa que um sistema de equações é resolvido a

cada passo de tempo. Em problemas com um elevado número de graus de liberdade, o FETD pode se tornar impraticável devido ao enorme custo computacional.

O DGTD é um método promissor na modelagem de problemas de multiescala. Existem várias versões deste método que podem ser encontradas na literatura. Uma versão com ordem de discretização espacial arbitrária é apresentada por [Hesthaven e Warburton \(2007\)](#). Neste trabalho foi demonstrado que o método é estável e converge com muita rapidez. Uma revisão do método Galerkin descontínuo, aplicado à nanofotônica pode ser encontrada em [Busch *et al.* \(2011\)](#).

O DG pode ser considerado como uma versão combinada do método de elementos finitos e o método de volumes finitos (FVM - *finite volume method*). Como o FEM, o DG usa malhas não estruturadas na discretização do domínio computacional. Isso proporciona uma melhor adaptação aos contornos e às várias escalas, trazendo uma maior precisão em relação aos métodos que usam *grids* estruturados, como é o caso do FDTD. Este ajuste evita o fenômeno de *staircase*. O fato dos elementos serem descontínuos implica em sistemas locais com matrizes pequenas. Quando são usados polinômios ortogonais como funções de base, não é necessário a inversão de matrizes. Como os elementos são descontínuos, os subsistemas locais podem ser resolvidos de forma paralela. Como no FVM, os elementos são descontínuos e a continuidade deve ser imposta pelo fluxo numérico. A maioria dos trabalhos encontrados na literatura usam o fluxo *upwind*, o fluxo central ou uma versão combinada dos dois.

Uma outra virtude do DGTD está na integração do tempo, que pode ser feita com métodos implícitos, explícitos ou um método híbrido (*implicit-explicit* - IMEX). Métodos de integração do tempo implícitos não são condicionalmente estáveis. Isto significa que a estabilidade destes esquemas não depende do passo de tempo. Por outro lado, nestes esquemas é necessário resolver um sistema de equações em cada instante de tempo. Estes métodos são normalmente mais complexos que os explícitos e mais difíceis de serem implementados. Além disso, uma grande disparidade entre as dimensões dos elementos da malha pode levar a problemas de condicionamento nos esquemas implícitos ([Angulo *et al.* \(2014\)](#)).

Os métodos explícitos são simples e diretos. Nestes esquemas não é necessário a inversão de matrizes. Geralmente, isso os torna muito mais eficientes do ponto de vista do tempo computacional. Porém, os métodos explícitos são condicionalmente estáveis. Dessa forma, o passo de tempo Δt está sujeito a uma condição CFL para que a estabilidade seja mantida. Em problemas de multiescala, onde o tamanho dos elementos é extremamente discrepante, o passo de tempo global deve respeitar o menor passo de tempo local, que é imposto pelo menor elemento. Isto pode levar a um esforço computacional muito grande. Os métodos mais utilizados em conjunto com o DG são: o Runge Kutta e o *leapfrog*. Todavia, outros algoritmos também podem ser usados, como Lax-Wendroff ([Alvarez *et al.* \(2015\)](#); [Hesthaven e Warburton \(2002b\)](#); [Yang *et al.* \(2013\)](#)). Todos estes, são esquemas explícitos.

É visto que o DGTD apresenta um potencial muito grande para o tratamento de problemas de nanofotônica que apresentam multiescala. Isto se deve à sua capacidade de usar funções polinomiais de alta ordem. Essa característica é essencial no combate à dispersão e à dissipação numérica. Este método é naturalmente paralelizável, trazendo a possibilidade de diminuir significativamente o tempo das simulações. O uso de malha não estruturada o permite representar precisamente partes arredondadas do domínio e o fato de ser formado por elementos descontínuos o torna apto ao refinamento adaptativo tipo *hp* local. Dessa forma, é possível refinar a malha somente onde é realmente necessário.

A integração do tempo pode se tornar um limitador quando se trata do tempo computacional, pois tanto os métodos implícitos quanto os explícitos apresentam sérias limitações quando aplicados a problemas com múltiplas escalas. Para superar estas limitações pesquisadores adotam basicamente dois caminhos: usar métodos híbridos, IMEX, ou usar métodos explícitos com passos de tempo locais (*local time stepping* - LTS). Uma vantagem dos esquemas LTS em relação a estratégias implícitas é que o primeiro pode ser usado recursivamente e é facilmente paralelizável. Além disso, tamanhos de elementos de malha muito díspares podem levar a problemas de condicionamento em métodos implícitos [Angulo et al. \(2014\)](#). Por isso, os esquemas LTS são estratégias mais comumente usadas.

1.2.2.1 LTS-DGTD

Muitos métodos LTS têm sido propostos com o objetivo de melhorar o desempenho do DGTD em termos de custo computacional. Estes métodos visam aumentar o passo de tempo global, mantendo a estabilidade do esquema. Os primeiros métodos que surgiram eram baseados no *leapfrog*, mas recentemente também encontramos métodos LTS baseados no Runge-Kutta. Um importante aspecto de um método LTS é a forma com que é satisfeito o fluxo numérico. Por exemplo, considere um elemento k cujo passo de tempo máximo é dado por Δt^k e um elemento vizinho m com passo de tempo Δt^m . Considere também que $\Delta t^k < \Delta t^m$. Supondo que no elemento k os campos são avaliados com o intervalo Δt^k e no elemento m com Δt^m , para que o fluxo numérico seja satisfeito, é necessário que o elemento k possua os valores dos campos em m avaliados com intervalo Δt^k . Os métodos encontrados na literatura satisfazem essa condição de diferentes maneiras. A seguir, serão discutidos alguns aspectos dos métodos mais importantes encontrados na literatura.

Em [Cohen et al. \(2006\)](#) é apresentado um método chamado de 2-classes (*2-class*) que é baseado no LF. Nesse método os elementos do domínio computacional são divididos em duas classes, de acordo com o passo de tempo máximo permitido em cada um deles. Na primeira classe, os elementos serão atualizados com um passo de tempo Δt . Já na segunda os elementos serão atualizados com um passo de tempo que é um múltiplo de Δt . Isto é feito para garantir

a coincidência em cada etapa do processo. Para satisfazer o fluxo numérico nos elementos da classe 1, que fazem fronteira com elementos da classe 2, são feitas interpolações com os valores de campo obtidos nos elementos da classe 2. Este método mostra eficiência quando a malha é formada com elementos muito pequenos e elementos muito grandes. Em malhas que os elementos possuem diversos tamanhos seria necessário o uso de mais classes. Porém o uso de interpolações aumenta o custo computacional e ainda insere características dissipativas ao esquema (Montseny *et al.* (2008)).

Para superar as limitações encontradas pelo método 2-classes, em Piperno (2006) é proposto um método multi-classe, de segunda ordem, que não faz uso de interpolações. Neste método é usado o esquema de Verlet que é uma reorganização do método LF em três etapas. Neste método, os elementos são divididos em N classes. A divisão é baseada em critérios geométricos e/ou físicos que determinam o maior passo de tempo em cada elemento. Considerando que o passo global seja Δt , para uma determinada classe κ os campos serão calculados com um intervalo dado por: $\Delta t^\kappa = \Delta t / 2^{N-\kappa}$. Dessa forma, os passos de tempo em cada classe são múltiplos de uma potência de 2 do menor passo de tempo pertencente a classe $\kappa = 1$. Elementos de classes mais baixas serão avaliados um número maior de vezes. Para satisfazer o fluxo numérico entre elementos vizinhos que residem em camadas diferentes não são feitas interpolações, em vez disso, são usados valores de campo calculados em um instante anterior. Apesar de evitar interpolações, o uso de valores de campo em um instante ligeiramente diferente pode inserir imprecisão ao esquema, principalmente em simulações onde os campos variam com frequências muito altas.

Um método chamado *leapfrog* recursivo é apresentado em Montseny *et al.* (2008). Este método é similar ao descrito anteriormente, com duas diferenças fundamentais. A primeira é que neste método é usado o LF2 no lugar do esquema de Verlet. São criados N_s estágios intermediários no LF2 de acordo com cada classe. A segunda diferença está no passo de tempo local. Considerando que o menor passo de tempo seja dado por Δt , o passo de tempo local em cada classe κ é dado por: $(2m+1)^{N-\kappa} \Delta t$, sendo m um inteiro positivo. O parâmetro m traz uma maior flexibilidade para os passos de tempo locais quando comparado ao método multi-classe. Porém, os passos de tempo locais, em cada classe $1 < \kappa \leq N$, devem ainda ser múltiplos de Δt . Uma outra limitação é que o maior passo de tempo que mantém a estabilidade do LF2 recursivo é em torno de 80% do passo de tempo para o LF2 clássico.

Em Diaz e Grote (2009) foi desenvolvido um estudo rigoroso sobre a estabilidade e dispersão de esquemas de alta ordem LF-LTS aplicados à equação de onda de segunda ordem. Este estudo foi feito por meio de uma análise de autovalores. Os autores mostram que o LTS introduz dispersão numérica e pode produzir instabilidades, se o passo de tempo global não é ligeiramente reduzido em relação a uma implementação clássica. Também é mostrado que a estabilidade global poderia ser melhorada através da ampliação da região que possui um menor

passo tempo.

O três métodos descritos até agora usam o *leapfrog* como base para o esquema LTS. Contudo, quando funções de base de alta ordem são utilizadas, é necessário que a integração do tempo também seja de alta ordem. Isso é necessário para manter a alta convergência do esquema totalmente discreto. O método de alta ordem mais utilizado em conjunto com o DG é o LSERK4 (*low storage explicit* Runge Kutta de quarta ordem). Também existem trabalhos que propõem métodos LTS baseados no LSERK4. Talvez o mais importante seja o apresentado em Angulo *et al.* (2014), chamado de CPLTS (*causal path local time stepping*). Este método LTS pode ser aplicado tanto ao LSERK4 quanto ao LF2, dando origem ao LSERK4-CPLTS e ao LF2-CPLTS, respectivamente.

No CPLTS, os elementos são divididos em camadas, de acordo com o passo de tempo local Δt^m . Assim, um elemento k pertencerá a uma camada m com $m = [0, \dots, N_m - 1]$. Os elementos são organizados nas camadas, de forma que: os que possuem Δt^m menores, são agrupados na camada, m , mais baixa. Assim, tem-se que: $\Delta t^m < \Delta t^{m+1}$. Em camadas mais baixas são realizadas um número maior de etapas onde os campos são atualizados. Elementos vizinhos que não pertencem a mesma camada devem pertencer a uma camada adjacente.

Para o cálculo do fluxo numérico entre elementos de camadas adjacentes, são realizadas etapas intermediárias nas camadas com maior Δt . Na versão baseada no LF2, isto é necessário apenas nos elementos vizinhos que pertencem a camada superior (primeira vizinhança). Já no LSERK, também é necessário calcular etapas intermediárias em elementos adjacentes aos da fronteira, para que o fluxo seja calculado. Neste método, os passos locais, em cada camada, podem ser escolhidos com maior liberdade que nos casos anteriores.

Quando aplicada ao LSERK4, a implementação CPLTS, em alguns casos, pode apresentar uma redução de metade do tempo gasto pelo método de Montseny. Já em outros casos pode não apresentar nenhuma melhoria no desempenho, mas a dissipação numérica é reduzida. Em ambos os casos, as propriedades dispersivas do esquema não parecem ser significativamente afetadas. Já o LF2-CPLTS, apresenta um desempenho pior por um fator de 2, quando comparado a técnica de Montseny. No entanto, esta técnica parece apresentar melhores propriedades dispersivas do que a abordagem do Montseny e tem melhor adaptabilidade a problemas de multiescala (Angulo *et al.* (2014)).

Apesar de apresentar bons resultados, o CPLTS apresenta limitações. O LF2-CPLTS não é estável quando usado com o fluxo central, sendo necessário o uso do fluxo *upwind*. Isso destrói a propriedade simplética do LF2. Apesar do LSERK4-CPLTS permitir que os intervalos Δt^m sejam variáveis em cada camada, é necessário que a camada superior tenha pelo menos 5/3 vezes o número de elementos na camada inferior, para que um ganho computacional seja real. Do contrário, é melhor que os estágios intermediários em camadas inferiores sejam uniformemente distribuídos, de forma que: $\Delta t^m = 1/N_s$. Uma grande limitação do LSERK4-CPLTS está

no fato desse esquema necessitar de pelo menos 4 subcamadas de vizinhos (quatro vizinhanças) entre classes adjacentes a fim de satisfazer o fluxo numérico. Isso se deve ao fato do LSERK4 ter 5 estágios. Essa característica limita o uso desse esquema em malhas onde essa restrição não possa ser satisfeita.

É visto que os métodos LTS-DGTD encontrados na literatura apresentam bons resultados. Porém, ainda podem ser melhorados, a fim de reduzir o tempo computacional e as características dissipativas do esquema. Além disso, em problemas onde o número de elementos extremamente pequenos é muito grande, estes métodos podem não apresentar grande melhoria em relação aos métodos clássicos. Dessa forma, o estudo de métodos LTS-DGTD aplicados à problemas com múltipla escalas é de grande interesse.

1.3 Contribuições da Tese

Neste trabalho a acurácia e a eficiência de formulações baseadas no método de Galerkin, na simulação de dispositivos nanofotônicos com dimensões que possuem múltiplas escalas são estudadas. O cenário para esse estudo, são estruturas baseadas em guias de ondas planares com base em um cristal fotônico bidimensional. Como na maioria dos casos encontrados na literatura, simulações 2D são realizadas. Acredita-se que os resultados que serão apresentados não têm perda de generalidade e podem ser estendidos para o caso 3D.

O objetivo geral desse trabalho é o estudo da eficiência do método multi classe, *mC-SSPMS*, em um esquema DGTD, na simulação da propagação em dispositivos baseados em cristal fotônico. Neste sentido:

- As estruturas de bandas de cristais fotônicos bidimensionais formados por uma rede quadrada de pilares dielétricos são modeladas e calculadas. Para isso, códigos FEM e NEM que foram desenvolvidos com base na modelagem que será descrita no Capítulo 2 são utilizados.
- É desenvolvida e validada uma estratégia multiclasse baseada no método de integração do tempo, SSPMS, que foi aplicada em conjunto com a discretização espacial DG, trazendo ganhos de eficiência computacional nas simulações que envolveram múltiplas escalas.
- São desenvolvidos códigos *mC-SSPMS*, validados usando um problema de referência. Em seguida, os códigos desenvolvidos são aplicados e validados na simulação da propagação eletromagnética nos dispositivos propostos.
- A estratégia proposta é comparada com a versão padrão e outros métodos encontrados na literatura.
- Uma forma de obter um conhecimento prévio acerca de quantas classes utilizar a fim de atingir um aumento da eficiência computacional é também desenvolvida.

Fundamentação Teórica

Esse capítulo traz os fundamentos dos cristais fotônicos e da fotônica computacional associada a estas estruturas. Aqui é demonstrado que: as bandas de frequências das ondas eletromagnéticas que podem propagar no interior dos cristais fotônicos podem ser calculadas a partir de um problema de autovalor, definido pela manipulação das equações de Maxwell e das relações constitutivas do meio. Nesse capítulo são abordadas as relações entre os modos eletromagnéticos e as diversas simetrias que pode ter um cristal fotônico bidimensional. São introduzidos conceitos fundamentais dos cristais bidimensionais formados por uma rede quadrada de colunas dielétricas. Também são discutidos os fundamentos de defeitos em cristais fotônicos e os conceitos de cavidades e guias. Por fim, são introduzidas noções sobre fotônica computacional, abordando as principais classes de problemas e os métodos mais usados.

2.1 Fundamentos de Cristais Fotônicos

Nessa seção serão discutidos os aspectos fundamentais dos cristais fotônicos. Para isso, a discussão será iniciada definindo um problema de autovalor a partir das equações de Maxwell. Em seguida, é mostrado que apenas observando suas simetrias, é possível entender o comportamento dos campos eletromagnéticos no interior do cristal. Na sequência são discutidos os aspectos fundamentais dos cristais 2D. No fim da seção são abordados os defeitos lineares que formam os guias de ondas.

2.1.1 Equações de Maxwell em Meios Dielétricos

A caracterização dos cristais fotônicos é basicamente o levantamento do diagrama de bandas. Dessa forma, é adequado trabalhar com as equações de Maxwell no domínio da frequência. O sistema considerado aqui, é livre de cargas e correntes. Considerando o meio dielétrico, tem-se que $\mu = \mu_0$ e $\varepsilon = \varepsilon_r \varepsilon_0$. Assim,

$$\nabla \cdot \mathbf{H}(\mathbf{r}) = 0 \quad (2.1)$$

$$\nabla \cdot (\varepsilon(\mathbf{r})\mathbf{E}(\mathbf{r})) = 0 \quad (2.2)$$

$$\nabla \times \mathbf{E}(\mathbf{r}) - j\omega\mu_0\mathbf{H}(\mathbf{r}) = 0 \quad (2.3)$$

$$\nabla \times \mathbf{H}(\mathbf{r}) + j\omega\varepsilon(\mathbf{r})\mathbf{E}(\mathbf{r}) = 0 \quad (2.4)$$

As equações (2.1) e (2.2) impõem a condição de transversalidade aos campos. Trabalhando as equações acima chega-se às expressões:

$$-\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) = \varepsilon_r(\mathbf{r}) \left(\frac{\omega}{c} \right)^2 \mathbf{E}(\mathbf{r}) \quad (2.5)$$

$$\nabla \times \frac{1}{\varepsilon_r(\mathbf{r})} \nabla \times \mathbf{H}(\mathbf{r}) = \left(\frac{\omega}{c} \right)^2 \mathbf{H}(\mathbf{r}) \quad (2.6)$$

que são as equações de onda, na forma vetorial, em função dos campos. Na modelagem de meios dielétricos mistos, é mais conveniente trabalhar com a equação envolvendo o campo magnético, pois o operador associado ao problema é Hermitiano e é mais simples garantir a condição de transversalidade, dada por (2.1), já que esta não envolve ε_r . De posse do campo magnético, é possível encontrar o campo elétrico da seguinte forma: $\mathbf{E}(\mathbf{r}) = j/\omega\varepsilon_0\varepsilon_r(\mathbf{r})\nabla \times \mathbf{H}(\mathbf{r})$ e assim a condição (2.2) também é garantida indiretamente. A equação (2.6) pode ser reescrita em termos de um operador diferencial, Θ , que representa o sistema (Joannopoulos *et al.* (2011)).

$$\Theta\mathbf{H}(\mathbf{r}) = \left(\frac{\omega}{c} \right)^2 \mathbf{H}(\mathbf{r}) \quad (2.7)$$

onde:

$$\Theta = \nabla \times \frac{1}{\varepsilon_r(\mathbf{r})} \nabla \times \quad (2.8)$$

É visto que (2.7) representa um problema de autovalor, onde o termo $(\omega/c)^2$ é o autovalor, o campo $\mathbf{H}$ é o autovetor e Θ é o operador associado. Este operador tem a característica importante

de ser Hermitiano, ou seja, $\Theta = \Theta^*$. Sendo Θ^* o transposto conjugado de Θ . Operadores Hermitianos possuem autovalores reais e positivos além de autovetores ortogonais (Arfken e Weber (1999)).

2.1.2 Simetrias relacionadas ao Cristal Fotônico

É possível ter um conhecimento prévio sobre os modos relacionados à uma estrutura, simplesmente observando sua simetria, ou seja, a simetria do sistema pode ajudar a entender suas propriedades eletromagnéticas. Ao longo deste texto será mostrado como os modos eletromagnéticos podem ser classificados levando em conta a simetria do sistema.

2.1.2.1 Inversão

Esse tipo de simetria implica em que o sistema seja invariante à operação de inversão em relação a um ponto, chamado ponto de inversão. Para um sistema invariante à inversão, que não possui degenerescência, os modos podem ser classificados como: par ou ímpar, ou seja, $\mathbf{H}(\mathbf{r}) = \mathbf{H}(-\mathbf{r})$ (modo par) ou $\mathbf{H}(\mathbf{r}) = -\mathbf{H}(-\mathbf{r})$ (modo ímpar). Se o sistema apresenta degenerescência, existirão modos com mesma frequência e perfis de campo que não diferem entre si apenas por uma constante multiplicativa. Contudo sempre é possível, através de uma combinação linear desses modos, construir modos que sejam ímpares ou pares (Joannopoulos *et al.* (2011)).

2.1.2.2 Simetria Translacional Contínua

Uma outra simetria de grande interesse que um sistema pode apresentar é a translacional contínua. Tal sistema é invariante à translação em uma dada direção. Um sistema que possui simetria translacional contínua em todas as três direções é um meio homogêneo, ou seja, $\epsilon(\mathbf{r})$ é uma constante. Um modo em um meio homogêneo deve ter a forma:

$$\mathbf{H}_{\mathbf{k}}(\mathbf{r}) = \mathbf{H}_0 e^{j\mathbf{k} \cdot \mathbf{r}} \quad (2.9)$$

onde $\mathbf{H}_0$ é um vetor constante que define a polarização da onda plana. Devendo ser respeitada a condição de transversalidade:

$$\nabla \cdot \mathbf{H}_{\mathbf{k}}(\mathbf{r}) = 0$$

$$\mathbf{k} \cdot \mathbf{H}_0 = 0 \quad (2.10)$$

É possível verificar que $\mathbf{H}_{\mathbf{k}}$ é solução da equação (2.6). Os autovalores são $(\omega/c)^2 = |\mathbf{k}|^2/\epsilon_r$, levando à relação de dispersão: $\omega = c|\mathbf{k}|/\sqrt{\epsilon_r}$ (Joannopoulos *et al.* (2011)). Dessa forma, ondas planas são classificadas pelo seu respectivo vetor de onda.

2.1.2.3 Simetria Translacional Discreta

Cristais fotônicos, assim como cristais tradicionais possuem simetria translacional discreta, ou seja, a estrutura é invariante a uma translação em uma distância que é um múltiplo de um certo comprimento (período). A Figura 2.1 traz um exemplo de um sistema com essa simetria. Esse sistema possui simetria translacional discreta na direção x . O período de repetição é chamado de constante de rede, a , e o vetor primitivo da rede (*primitive lattice vector*) é dado por: $\mathbf{a} = a\hat{\mathbf{x}}$. Consequentemente, $\epsilon(\mathbf{r}) = \epsilon(\mathbf{r} \pm \mathbf{a})$, ou ainda, $\epsilon(\mathbf{r}) = \epsilon(\mathbf{r} + \mathbf{R})$. Onde $\mathbf{R} = \ell\mathbf{a}$, sendo ℓ um inteiro. É possível representar esse sistema através de uma célula unitária, que é mostrada na Figura 2.1 pelo retângulo tracejado.

A fim de entender as implicações da simetria discreta nos modos eletromagnéticos, considere um operador translacional $T_{\mathbf{R}}$, tal que $T_{\mathbf{R}}\mathbf{H}_{\mathbf{k}}(\mathbf{r}) = \mathbf{H}_{\mathbf{k}}(\mathbf{r} + \mathbf{R})$. Por causa da simetria, $\mathbf{H}_{\mathbf{k}}$ deve ser uma autofunção do operador translacional discreto, com deslocamentos $\mathbf{R} = \ell\mathbf{a}$. Portanto, os modos esperados devem ser ondas planas, $e^{jk_x x}$ (Joannopoulos *et al.* (2011)). Portanto:

$$T_{\mathbf{R}}e^{jk_x x} = e^{jk_x(x-\ell a)} = e^{-jk_x \ell a} e^{jk_x x} \quad (2.11)$$

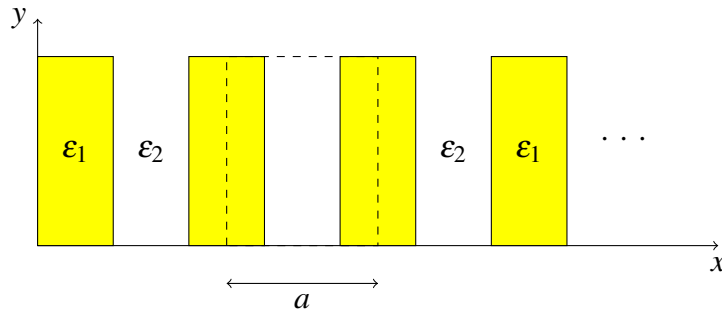


Figura 2.1: Simetria Translacional Discreta

Observando a equação acima nota-se que nem todos os valores diferentes de k_x levam a autovalores também diferentes. Considerando por exemplo um outro vetor de onda $k_x + 2\pi/a$:

$$\begin{aligned}
 T_{\mathbf{R}} e^{j(k_x + \frac{2\pi}{a})x} &= e^{-j\ell a(k_x + \frac{2\pi}{a})} e^{j(k_x + \frac{2\pi}{a})x} \\
 T_{\mathbf{R}} e^{j(k_x + \frac{2\pi}{a})x} &= e^{-j2\pi\ell} e^{-jk_x\ell a} e^{j(k_x + \frac{2\pi}{a})x} \\
 T_{\mathbf{R}} e^{j(k_x + \frac{2\pi}{a})x} &= e^{-jk_x\ell a} e^{j(k_x + \frac{2\pi}{a})x}
 \end{aligned} \tag{2.12}$$

O autovalor associado ao modo com a componente do vetor de onda $k_x + 2\pi/a$ é o mesmo do modo com vetor k_x . De fato, todos os modos com vetor de onda na forma $k_x + m(2\pi/a)$, onde m é um inteiro, formam um conjunto de modos degenerados. Para um operador $T_{\mathbf{R}}$, todos esses modos possuem o mesmo autovalor $e^{-jk_x\ell a}$. Somando $b = 2\pi/a$ a k_x os autovalores permanecem os mesmos. Assim, é possível definir um vetor de rede recíproca primitivo, $\mathbf{b} = b\hat{\mathbf{x}}$. A combinação linear de autofunções, também são autofunções com os mesmo autovalores. Então, tomando uma combinação linear dos modos originais e considerando que o sistema é infinito em y ,

$$\begin{aligned}
 \mathbf{H}_{k_x, k_y} &= e^{jk_y y} \sum_m \mathbf{c}_{k_x, m} e^{j(k_x + mb)x} \\
 \mathbf{H}_{k_x, k_y} &= e^{jk_x x} e^{jk_y y} \sum_m \mathbf{c}_{k_x, m}(x) e^{jmbx} \\
 \mathbf{H}_{k_x, k_y} &= e^{jk_x x} e^{jk_y y} \mathbf{u}_{k_x}(x)
 \end{aligned} \tag{2.13}$$

Na equação (2.13) os $\mathbf{c}$'s são os coeficientes da série. Pode ser visto também que $\mathbf{u}_{k_x}(x)$ é periódica em x , de forma que: $\mathbf{u}(x + \ell a) = \mathbf{u}(x)$.

A simetria discreta impõe que a dependência dos modos $\mathbf{H}$ com x seja o produto de uma onda plana com uma função periódica em x . Isto pode ser pensado como uma onda plana como no espaço livre mas modulada pela periodicidade da rede:

$$\mathbf{H}(\dots, x, \dots) \propto e^{jk_x x} \mathbf{u}_{k_x}(\dots, x, \dots) \tag{2.14}$$

Esse resultado é geralmente conhecido como teorema de Bloch (Kittel *et al.* (1976)). $\mathbf{H}$ é chamado na Física do Estado Sólido de estado de Bloch, e na Mecânica de modo de Floquet. Uma importante característica dos estados de Bloch é que um estado com vetor de onda k_x é idêntico a outro com vetor $k_x + mb$. Assim as frequências associadas a estes modos também devem ser periódicas, de forma que para k_x tem-se que: $\omega(k_x) = \omega(k_x + mb)$ (Joannopoulos *et al.* (2011)). Por este fato só é necessário considerar os vetores que estão na faixa: $-\pi/a < k_x \leq \pi/a$, dita primeira zona de Brilloin.

Analogamente, se o sistema possui simetria translacional discreta nas três dimensões, de acordo com o teorema de Bloch os modos devem ter a forma:

$$\mathbf{H}_{\mathbf{k}} = e^{j\mathbf{k} \cdot \mathbf{r}} \mathbf{u}_{\mathbf{k}}(\mathbf{r}) \quad (2.15)$$

Os modos são classificados de acordo com o vetor de onda de Bloch $\mathbf{k} = k_1 \mathbf{b}_1 + k_2 \mathbf{b}_2 + k_3 \mathbf{b}_3$, que pertence a primeira zona de Brillouin. A função $\mathbf{u}_{\mathbf{k}}$ é periódica: $\mathbf{u}_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) = \mathbf{u}_{\mathbf{k}}(\mathbf{r})$, para um vetor de rede $\mathbf{R} = \ell \mathbf{a}_1 + m \mathbf{a}_2 + n \mathbf{a}_3$, sendo $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi \delta_{ij}$.

2.1.2.4 Estrutura de Bandas Fotônicas

Na sub-seção anterior foi demonstrado que, os modos eletromagnéticos de cristais fotônicos que possuem simetria translacional discreta nas três dimensões podem ser escritos como modos de Bloch. Substituindo um modo de Bloch na equação 2.7 tem-se que:

$$\begin{aligned} \Theta \mathbf{H}_{\mathbf{k}} &= \left(\frac{\omega(\mathbf{k})}{c} \right)^2 \mathbf{H}_{\mathbf{k}} \\ \nabla \times \frac{1}{\varepsilon(\mathbf{r})} \nabla \times \mathbf{H}_{\mathbf{k}} &= \left(\frac{\omega(\mathbf{k})}{c} \right)^2 \mathbf{H}_{\mathbf{k}} \\ \nabla \times \frac{1}{\varepsilon(\mathbf{r})} \nabla \times e^{j\mathbf{k} \cdot \mathbf{r}} \mathbf{u}_{\mathbf{k}}(\mathbf{r}) &= \left(\frac{\omega(\mathbf{k})}{c} \right)^2 e^{j\mathbf{k} \cdot \mathbf{r}} \mathbf{u}_{\mathbf{k}}(\mathbf{r}) \\ (j\mathbf{k} + \nabla) \times \frac{1}{\varepsilon(\mathbf{r})} (j\mathbf{k} + \nabla) \times \mathbf{u}_{\mathbf{k}}(\mathbf{r}) &= \left(\frac{\omega(\mathbf{k})}{c} \right)^2 \mathbf{u}_{\mathbf{k}}(\mathbf{r}) \end{aligned} \quad (2.16)$$

A equação (2.16) pode ser reescrita como:

$$\Theta_{\mathbf{k}} \mathbf{u}_{\mathbf{k}}(\mathbf{r}) = \left(\frac{\omega(\mathbf{k})}{c} \right)^2 \mathbf{u}_{\mathbf{k}}(\mathbf{r}) \quad (2.17)$$

$$\Theta_{\mathbf{k}} = (j\mathbf{k} + \nabla) \times \frac{1}{\varepsilon(\mathbf{r})} (j\mathbf{k} + \nabla) \times \quad (2.18)$$

O novo operador (2.18) também é Hermitiano e depende de $\mathbf{k}$. A função $\mathbf{u}_{\mathbf{k}}$ e consequentemente os modos $\mathbf{H}_{\mathbf{k}}$, são obtidos resolvendo o problema de autovalor dado por (2.17) e sujeito as condições de transversalidade e periodicidade (Joannopoulos *et al.* (2011)):

$$(j\mathbf{k} + \nabla) \cdot \mathbf{u}_{\mathbf{k}}(\mathbf{r}) = 0 \quad (\text{transversalidade}) \quad (2.19)$$

$$\mathbf{u}_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) = \mathbf{u}_{\mathbf{k}}(\mathbf{r}) \quad (\text{periodicidade}) \quad (2.20)$$

As condições periódicas nos permitem resolver o problema em uma célula unitária do cristal fotônico. É esperado para um dado $\mathbf{k}$, um conjunto infinito de modos com faixas de frequência discretas, $\omega_n(\mathbf{k})$ (n define o modo). $\omega_n(\mathbf{k})$ são funções que variam continuamente com $\mathbf{k}$. O conjunto de todas as $\omega_n(\mathbf{k})$ formam o diagrama de estrutura de bandas do cristal.

2.1.2.5 Simetria Rotacional e Zona Irredutível de Brillouin

Se o sistema possui simetria rotacional, as frequências $\omega_n(\mathbf{k})$ possuem redundância dentro da zona de Brillouin. Se o cristal fotônico possui simetria rotacional e reflexão especular ou simetria inversa as frequências $\omega_n(\mathbf{k})$ também possuirão. Essa coleção de simetrias é conhecida como ponto de grupo do cristal (*point group*). Adicionalmente, se $\omega_n(\mathbf{k})$ possui todas as simetrias citadas, só não haverá redundância na zona irredutível de Brillouin (Joannopoulos *et al.* (2011)).

2.1.2.6 Simetria Especular e Separação de Modos

A Simetria especular é muito importante, pois, permite separar os modos eletromagnéticos em duas polarizações: TE ou TM (Joannopoulos *et al.* (2011)). Isto é conveniente, pois fornece informações sobre a simetria dos modos e simplifica a solução do problema de autovalor. Observando a estrutura da Figura 2.1, percebe-se que ela possui simetria especular em relação ao eixo x . Logo, considerando qualquer plano yz apenas duas situações são possíveis:

- $\mathbf{H}_{\mathbf{k}}$ é perpendicular ao plano, enquanto que $\mathbf{E}_{\mathbf{k}}$ é paralelo.
- $\mathbf{E}_{\mathbf{k}}$ é perpendicular ao plano, enquanto que $\mathbf{H}_{\mathbf{k}}$ é paralelo.

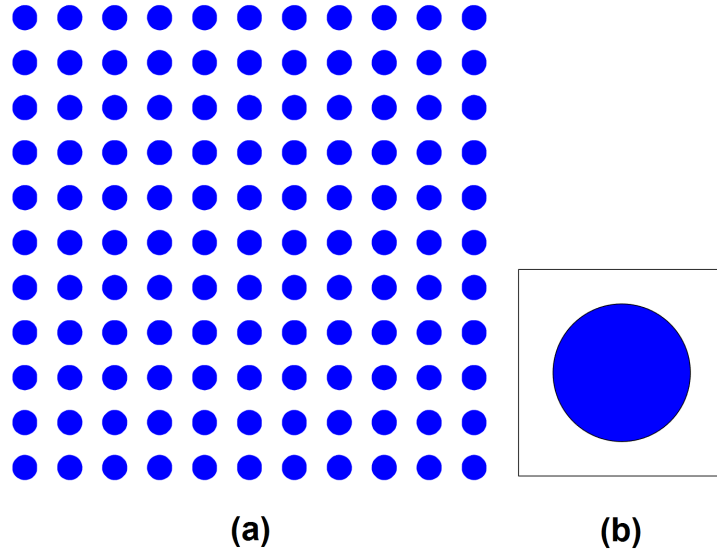


Figura 2.2: Cristais Fotônicos Bidimensionais

2.1.3 Cristal Fotônico Bidimensional

Agora serão introduzidos, de forma geral, conceitos associados aos cristais fotônicos bidimensionais formados por uma rede com simetria quadrada de colunas dielétricas, imersas em ar. Um exemplo desse cristal fotônico bidimensional é mostrado em perspectiva 2D na Figura 2.2. Em (a) é mostrada uma rede com simetria quadrada de colunas dielétricas infinitas, imersas em ar. Em (b) é mostrada sua célula unitária.

Para o estudo das bandas de frequência dos cristais 2D, é desenvolvida uma formulação variacional fraca, do problema de autovalor, definido a partir das equações de Maxwell. A implementação foi feita para duas formas de discretização: FEM e NEM. Todo esse desenvolvimento será mostrado mais a frente nesse capítulo. O diagrama que será apresentado nessa seção foi obtido usando uma implementação do método dos elementos finitos.

2.1.3.1 Modos de Bloch Bidimensionais

Como definido anteriormente, é possível entender o comportamento dos modos apenas observando a simetria do sistema. Como o sistema possui simetria translacional contínua em z , espera-se que os campos oscilem nesta direção como uma onda plana e não existirá restrição a k_z . No plano xy , o sistema possui simetria translacional discreta, e por isso, k_x e k_y devem pertencer a zona de Brillouin, e os modos devem oscilar como ondas de Bloch. Com essas considerações em mente, os modos estendidos ao interior desse cristal devem ter a seguinte forma:

$$\mathbf{H}_{(n,k_z,\mathbf{k}_{\parallel})}(\mathbf{r}) = e^{j\mathbf{k}_{\parallel} \cdot \boldsymbol{\rho}} e^{jk_z z} \mathbf{u}_{(n,k_z,\mathbf{k}_{\parallel})}(\boldsymbol{\rho}) \quad (2.21)$$

Na equação (2.21) tem-se que: $\boldsymbol{\rho} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}}$, $\mathbf{k}_{\parallel} = k_x\hat{\mathbf{x}} + k_y\hat{\mathbf{y}}$ e $\mathbf{u}(\boldsymbol{\rho} + \mathbf{R}) = \mathbf{u}(\boldsymbol{\rho})$, onde $\mathbf{R} = \ell_1 a\hat{\mathbf{x}} + \ell_2 a\hat{\mathbf{y}}$. O sistema possui simetria especular em relação ao plano xy . Consequentemente, os modos podem ser classificados em TE (E_x, E_y, H_z) e TM (H_x, H_y, E_z).

2.1.3.2 Rede Quadrada de Colunas Dielétricas

Por simplicidade, só serão considerados modos com $k_z = 0$. A Figura 2.3 traz o diagrama de estrutura de bandas de um cristal como o da figura 2.2, consistindo de colunas dielétricas com $\varepsilon_r = 8.9$ e $r/a = 0.2$. As curvas foram calculadas com o FEM clássico utilizando condições periódicas. O eixo vertical representa $\omega a/2\pi c$, que é a frequência normalizada. O eixo horizontal representa os vetores de onda de Bloch normalizados, $\mathbf{k}_{\parallel}$. Como esta estrutura tem simetria rotacional de 90° , para evitar redundância, só são considerados os vetores na zona irredutível de Brillouin. Os modos são plotados seguindo a sequência $\Gamma \rightarrow X \rightarrow M \rightarrow \Gamma$, onde: $\Gamma = (0,0)$, $X = (1/2,0)$ e $M = (1/2,1/2)$. O máximo e o mínimo de uma banda quase sempre ocorrem nas arestas da zona irredutível (Joannopoulos *et al.* (2011)). Pela Figura 2.3, nota-se que existe um *gap* de 38% (*gap-midgap*) para o modo TM, mas não para o modo TE. A explicação para este fato pode ser tirada do teorema variacional:

$$\mathcal{U}_f(\mathbf{H}) = \frac{(\Theta \mathbf{H}, \mathbf{H})}{(\mathbf{H}, \mathbf{H})} \quad (2.22)$$

Esse funcional, que é também conhecido como coeficiente de Rayleigh, é minimizado pelo menor autovalor de Θ , $(\omega_0/c)^2$. O funcional dado por (2.22) pode ser escrito em função do campo elétrico (Joannopoulos *et al.* (2011)):

$$\mathcal{U}_f(\mathbf{H}) = \frac{\int |\nabla \times \mathbf{E}(\mathbf{r})|^2 dV}{\int \varepsilon_r(\mathbf{r}) |\mathbf{E}(\mathbf{r})|^2 dV} \quad (2.23)$$

O denominador de (2.23) é a energia de campo elétrico da célula unitária. Logo o funcional deve atingir menores valores quanto maior for a energia. Isso é favorecido pela maior concentração do campo elétrico na região de maior constante dielétrica. Consequentemente, os modos de baixa frequência tendem a concentrar a maior parte de sua energia nas regiões de maior constante dielétrica. Somado a isso, o fato dos modos serem ortogonais, tem-se que: os modos de maior frequência tendem a concentrar grande parte de sua energia nas regiões de menor constante dielétrica. Portanto, é esperado que uma grande diferença na concentração da energia, na região de maior constante dielétrica, entre modos de bandas adjacentes, provoque

uma descontinuidade no espectro de frequência.

É importante notar que, para que haja uma grande diferença na concentração da energia, é necessário que a geometria do cristal permita a concentração do campo elétrico e que a diferença entre as permissividades elétricas dos materiais que compõem o cristal seja grande o bastante. No caso particular abordado aqui, para o modo TM, $\mathbf{E} = (0, 0, E_z)$. Isso torna mais fácil que o campo elétrico se concentre nas regiões de maior constante dielétrica, ou seja, nas colunas. Para o modo TE, o campo $\mathbf{E}$ está no plano xy , dificultando o acúmulo de energia nas regiões de maior constante dielétrica. Este contraste pode ser medido quantitativamente pelo fator de concentração, f_c , dado pela equação:

$$f_c = \frac{\int_{\Omega_\epsilon} \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r})|^2 dA}{\int_{\Omega} \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r})|^2 dA} \quad (2.24)$$

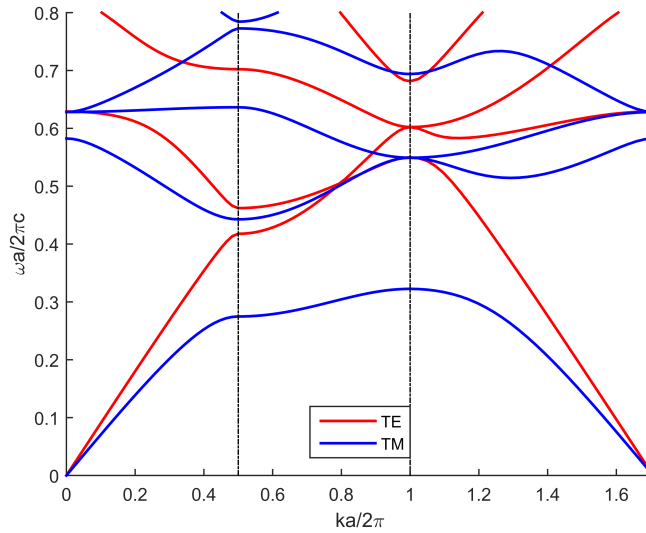


Figura 2.3: Estrutura de bandas de um cristal bidimensional de rede quadrada de colunas dielétricas com constante $\epsilon_r = 8.9$ e raio $r/a = 0.2$.

onde Ω_ϵ representa a região de maior constante dielétrica e Ω representa toda a célula unitária. O termo, dA , é o diferencial de área, em coordenadas cartesianas: $dA = dx dy$.

2.1.3.3 Defeitos Lineares e Guias de Ondas

É possível projetar e construir diversas aplicações de engenharia utilizando cristais fotônicos. Para isso, são inseridos defeitos na rede do cristal. Inserindo defeitos pontuais, cavidades ressonantes são criadas. No cristal ilustrado pela Figura 2.2, defeitos pontuais podem ser criados retirando ou alterando as características de uma coluna dielétrica. Inserindo defeitos lineares, são criados guias de ondas, que podem ser a base de outros dispositivos como biosensores

(Skivesen *et al.* (2007)), portas lógicas (Fu *et al.* (2013)), multiplexadores/demultiplexadores (Koshiba (2001)), e muitos outros encontrados na literatura. Um defeito linear pode ser inserido retirando ou alterando as características de uma linha de colunas dielétricas.

Pra construir um guia de ondas chamado W1, é retirada uma linha de colunas na direção x , como mostra a Figura 2.4. Na direção x a simetria translacional discreta é conservada, por isso os modos podem ser classificados em função de k_x . Como a simetria translacional contínua é mantida em z , apenas a propagação no plano xy pode ser considerada. Também serão considerados apenas os modos TM. A Figura 2.5 traz o diagrama de bandas projetado de um guia como o da Figura 2.4, tendo ω em função de k_x .

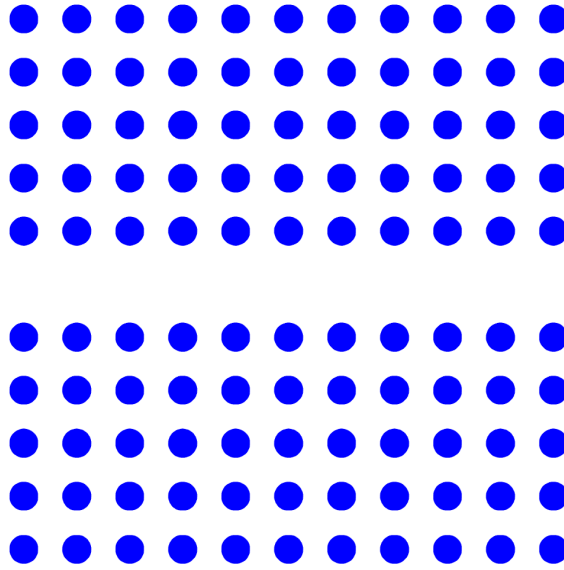


Figura 2.4: Guia W1 formado por uma linha de defeitos em um cristal fotônico bidimensional com rede quadrada de colunas dielétricas imersas no ar.

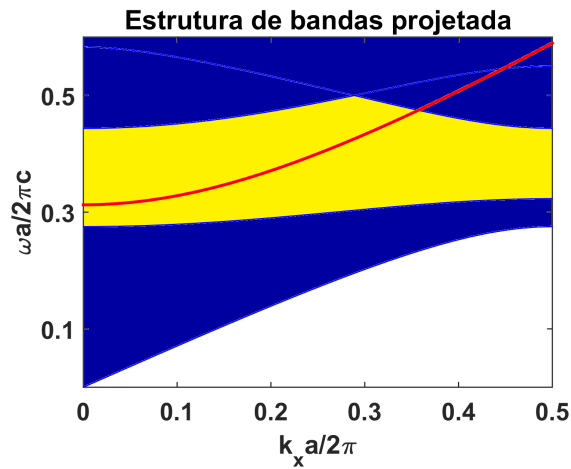


Figura 2.5: Estrutura de bandas projetada e diagrama de dispersão de um modo guiado em uma linha de defeitos de largura d .

Como representado na Figura 2.5, a linha de defeitos gera um único modo dentro do *gap* TM. Este modo é evanescente no interior do cristal. Existe uma diferença marcante entre modos localizados em defeitos pontuais e modos guiados por defeitos lineares. Para um modo ser confinado em um defeito pontual, sua frequência deve pertencer ao *gap*. Já um modo guiado deve ser considerado como função da frequência e do vetor de onda k_x . Assim, para um modo guiado é apenas necessário que a combinação (k_x, ω_0) não seja permitida no cristal.

Para testar a existência de um modo guiado no cristal, escolhe-se um específico par (k_x, ω_0) e verifica-se que para algum k_x este modo pertence a alguma banda. O processo de escolher um k_x e buscar as frequências para todos k_y é chamado de projeto da estrutura de bandas. O resultado desse processo é ilustrado pela região azul na Figura 2.5. Mantendo a simetria translacional discreta em uma direção, o número de formas possíveis de defeitos é muito grande. Retirando apenas uma linha de colunas, é criado um guia monomodo. Se são retiradas mais de uma, um guia multimodo é criado (Joannopoulos *et al.* (2011)).

2.1.3.4 Projetando Guias de Cristais 2D

A fim de ilustrar o projeto de um guia W1, considere o guia da Figura 2.4. As colunas têm raio igual a $0,2a$ e constante dielétrica $\epsilon_r = 8,9$. O *gap-midgap* dessa estrutura é $37.9321\% \approx 38\%$. O comprimento de onda escolhido foi $\lambda_0 = 1,5\mu m$ e a frequência $\omega a/2\pi c = 0,38$. Assim tem-se que: $a/\lambda_0 = 0,38$, $a = 0,57\mu m$, e $r = 0,114\mu m$. A Figura 2.5 traz a estrutura de bandas projetada para este guia com o FEM. O defeito tem largura $d = 2(a - r)$.

O diagrama de dispersão do modo fundamental é representado pela curva vermelha na Figura 2.5. As retas tracejadas representam os limites do *gap* completo e a região amarela representa as frequências dos modos que podem ser guiados. A região azul representa os modos estendidos no cristal, para k_x e k_y na zona irredutível de Brillouin. O diagrama de dispersão do modo guiado é dado por:

$$\frac{\omega(k_x)a}{2\pi c} = \sqrt{\left(\frac{k_x a}{2\pi}\right)^2 + \left(\frac{k_y a}{2\pi}\right)^2} \quad (2.25)$$

ou ainda:

$$\omega_n(k_{nx}) = \sqrt{k_{nx}^2 + k_{ny}^2} \quad (2.26)$$

Sabendo que $k_y = 2\pi/\lambda_y$, e que para o modo fundamental $\lambda_y = 2d = 4(a - r)$, $k_{ny} = a/4(a - r) = 0,3125$, e portanto:

$$\omega_n(k_x) = \sqrt{k_{nx}^2 + 0,3125^2} \quad (2.27)$$

A Figura 2.6 mostra a simulação do modo propagando no guia projetado.

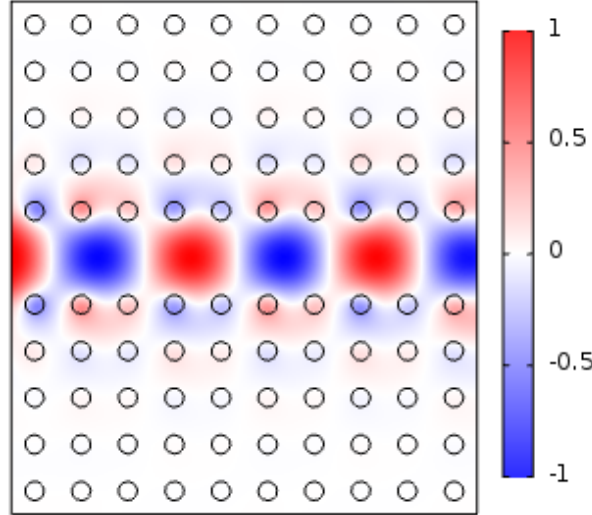


Figura 2.6: Modo TM propagando em um guia W1, formado por uma linha de defeitos com largura $d = 0,912\mu m$.

2.2 Modelagem das Bandas de Cristais Fotônicos 2D pelo Método de Galerkin

Essa seção descreve a modelagem da estrutura de bandas de um cristal fotônico bidimensional com rede de simetria quadrada, contudo a modelagem para cristais 2D e redes com simetrias diferentes são similares e podem ser desenvolvidas de forma análoga. As bandas de frequência e os perfis de campo dos modos do cristal são modelados como um problema de autovalor. Para isso, primeiro é determinada a forma forte e a partir dela, é desenvolvida a forma fraca. A forma fraca é discretizada pelo método de Galerkin. Por fim, as condições periódicas devem ser impostas.

2.2.1 Forma Forte

Para que a forma forte seja determinada é necessário definir a equação governante e as condições de fronteira.

2.2.1.1 Equação governante

Por causa da simetria especular, os campos são desacoplados, o que permite lidar com os modos TE e TM separadamente. Considerando o meio sem perdas, livre de correntes e cargas. Sendo os campos harmônicos, das equações de Maxwell e das relações constitutivas é possível derivar:

$$-\nabla \cdot \frac{1}{\epsilon_r(\mathbf{r})} \nabla H_z(\mathbf{r}) = \left(\frac{\omega}{c}\right)^2 H_z(\mathbf{r}) \quad (2.28)$$

$$-\nabla^2 E_z(\mathbf{r}) = \epsilon_r(\mathbf{r}) \left(\frac{\omega}{c}\right)^2 E_z(\mathbf{r}) \quad (2.29)$$

que são as equações de Helmholtz na forma escalar para os modos TE e TM respectivamente. Nas equações (2.28) e (2.29), E_z e H_z são componentes escalares dos campos elétrico e magnético. Em ambas as situações, os campos são orientados ao longo da direção z e variam em função de x e y . Isso garante a condição de ortogonalidade e evita o aparecimento de modos espúrios com frequência nula.

2.2.1.2 Condições Quasi-Periódicas

A Figura 2.7 traz a representação topológica de uma célula unitária com simetria quadrada, de lado a . Então, sendo a a constante de rede, o vetor de rede é dado por: $\mathbf{R} = a\hat{x} + a\hat{y}$. Na figura também é possível identificar o contorno: $\Gamma_b \cup \Gamma_r \cup \Gamma_u \cup \Gamma_l$. A célula unitária representa a menor porção da estrutura periódica, de forma que o perfil de campo nela é repetido por toda a estrutura do cristal. Logo, não é necessário resolver o problema em todo domínio, reduzindo significativamente o custo computacional.

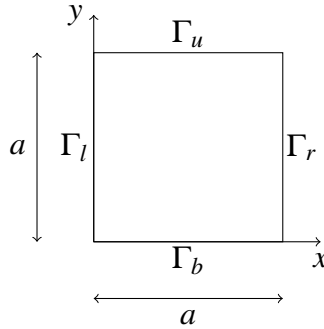


Figura 2.7: Representação topológica das fronteiras da célula unitária.

Pelo teorema de Floquet-Bloch, os campos que se propagam por um cristal fotônico possuem comportamento quasi-periódico e podem ser representados na forma:

$$U(\mathbf{r}) = u(\mathbf{r})e^{j\mathbf{k}\cdot\mathbf{r}} \quad (2.30)$$

onde $\mathbf{k} = k_x\hat{\mathbf{x}} + k_y\hat{\mathbf{y}}$ é o vetor de onda de Bloch. Sendo $u(\mathbf{r})$ periódica em $\mathbf{R}$:

$$u(\mathbf{r}) = u(\mathbf{r} + \mathbf{R}). \quad (2.31)$$

Utilizando o teorema de Floquet-Bloch são derivadas as condições de contorno para o problema. Estas são chamadas de condições quasi-periódicas e são dadas por (Schmidt e Kauf (2009)):

$$U(\mathbf{r} + a\hat{\mathbf{x}}) = U(\mathbf{r})e^{jk_x a} \quad (2.32)$$

$$U(\mathbf{r} + a\hat{\mathbf{y}}) = U(\mathbf{r})e^{jk_y a} \quad (2.33)$$

$$\frac{\partial U(\mathbf{r} + a\hat{\mathbf{x}})}{\partial n} = -\frac{\partial U(\mathbf{r})}{\partial n}e^{jk_x a} \quad (2.34)$$

$$\frac{\partial U(\mathbf{r} + a\hat{\mathbf{y}})}{\partial n} = -\frac{\partial U(\mathbf{r})}{\partial n}e^{jk_y a} \quad (2.35)$$

de posse das equações e das condições de fronteira, a formulação forte do problema proposto fica totalmente definida.

2.2.2 Forma Fraca

Para utilizar o FEM ou o NEM na solução do problema, é necessário encontrar a forma fraca. Assim, nesta sessão, é desenvolvida a forma fraca para os modos TE e TM.

2.2.2.1 Modo TE

A forma forte para o modo TE é dada pela equação (2.28), juntamente com as condições quasi-periódicas (2.32,...,2.35). Aqui a forma fraca é obtida pelo método dos resíduos ponderados. Considerando $V(\mathbf{r}) = v(\mathbf{r})e^{j\mathbf{k}\cdot\mathbf{r}}$, ou seja, que essa função tem as mesmas características do campo procurado, onde $v \in W$ e $W = \{v(\mathbf{r}) \in H^1; v(\mathbf{r} + \mathbf{R}) = v(\mathbf{r})\}$. Pelo método dos resíduos ponderados tem-se que:

$$-\int_{\Omega} (\nabla \cdot \frac{1}{\epsilon_r} \nabla H_z) V^* d\Omega = \int_{\Omega} (\frac{\omega}{c})^2 H_z V^* d\Omega \quad (2.36)$$

a aplicação do teorema da divergência, e um pouco de álgebra leva à:

$$\int_{\Omega} (\nabla \cdot (V^* \frac{1}{\epsilon_r} \nabla H_z)) d\Omega = \int_{\Gamma_{\Omega}} (V^* \frac{1}{\epsilon_r} \nabla H_z) \cdot \hat{\mathbf{n}} d\Gamma = \int_{\Gamma_{\Omega}} V^* \frac{1}{\epsilon_r} \frac{\partial H_z}{\partial n} d\Gamma. \quad (2.37)$$

onde $\hat{\mathbf{n}}$ é o vetor unitário normal à fronteira. Sabendo que $V^* = v(\mathbf{r})e^{-j\mathbf{k} \cdot \mathbf{r}}$ e que $v(\mathbf{r} + \mathbf{R}) = v(\mathbf{r})$, é fácil mostrar que $V^*(\mathbf{r} + a\hat{\mathbf{x}}) = V^*(\mathbf{r})e^{-jk_x a}$ e $V^*(\mathbf{r} + a\hat{\mathbf{y}}) = V^*(\mathbf{r})e^{-jk_y a}$. Assim, a integral de contorno dada por (2.37) se anula em função das condições de contorno e a equação (2.36) se reduz a:

$$\int_{\Omega} \frac{1}{\epsilon_r} \nabla H_z \cdot \nabla V^* d\Omega = (\frac{\omega}{c})^2 \int_{\Omega} H_z V^* d\Omega \quad (2.38)$$

como $H_z = ue^{j\mathbf{k} \cdot \mathbf{r}}$ e $V^* = ve^{-j\mathbf{k} \cdot \mathbf{r}}$, a forma fraca para o modo TE é dada por:

$$\int_{\Omega} \frac{1}{\epsilon_r} (\nabla u + j\mathbf{k}u) \cdot (\nabla v - j\mathbf{k}v) d\Omega = (\frac{\omega}{c})^2 \int_{\Omega} uvd\Omega \quad (2.39)$$

2.2.2.2 Modo TM

No modo TM, a equação governante é dada por (2.29) e as condições de fronteira são dadas pelas equações (2.32,...,2.35). Seguindo os mesmos procedimentos apresentados na subseção anterior, chega-se a forma fraca para o modo TM:

$$\int_{\Omega} (\nabla u + j\mathbf{k}u) \cdot (\nabla v - j\mathbf{k}v) d\Omega = (\frac{\omega}{c})^2 \int_{\Omega} \epsilon_r uvd\Omega \quad (2.40)$$

2.2.3 Discretização

A discretização é feita utilizando o método de Galerkin, ou seja:

$$\begin{aligned} u(x, y) &= \sum_i d_i \phi_i(x, y) \\ v(x, y) &= \sum_j c_j \phi_j(x, y) \end{aligned} \quad (2.41)$$

onde as funções ϕ são os interpolantes do FEM ou do NEM, dependendo de qual método se deseja usar. Quando as funções u e v são substituídas em (2.39) ou (2.40), o resultado é um problema de autovalor, que pode ser representado na forma matricial:

$$(\mathbf{A}(\mathbf{k}) - \lambda \mathbf{B})\mathbf{d} = 0. \quad (2.42)$$

As matrizes $\mathbf{A}$ e $\mathbf{B}$ mudam de acordo com (2.39) e (2.40). Para gerar o diagrama de estrutura de bandas do cristal, esse sistema deve ser resolvido para cada $\mathbf{k}$ discretizado nas arestas da zona irreduzível de Brillouin. A zona irreduzível de Brillouin é a região do espaço recíproco onde as frequências $\omega(\mathbf{k})$ e os perfis de campo $\mathbf{E}_{\mathbf{k}}$ e $\mathbf{H}_{\mathbf{k}}$, associados aos valores de $\mathbf{k}$ contidos nela, são repetidos por todo o espaço. Esta região é dada pelo triângulo com vértices $\Gamma = (0, 0)$, $X = (0, \pi/a)$ e $M = (\pi/a, \pi/a)$.

2.2.4 Funções de Base

Nesta seção são apresentadas as funções de base do método NEM. O método FEM não será abordado aqui porque já é bem conhecido e portanto, não se faz necessário. Ao leitor que não é familiarizado com o FEM, uma boa introdução deste método aplicado a problemas eletromagnéticos é encontrada em Polycarpou (2005).

2.2.4.1 Funções NEM

Os *natural neighbor coordinates*, como são chamadas as funções de forma do NEM, são construídas a partir do diagrama de Voronoi e da triangulação de Delaunay.

2.2.4.2 Diagrama de Voronoi

O diagrama de Voronoi é a decomposição de um domínio em subdomínios. Cada subdomínio é chamado de célula de Voronoi e é associado a um nó contido no domínio. Seja Ω uma região no espaço Euclidiano, onde estão distribuídos um conjunto de M nós. A célula de Voronoi T_I , associada a um nó x_I em Ω é dada por (Sukumar *et al.* (2001)):

$$T_I = \{x \in \mathbb{R} : \|x - x_I\| < \|x - x_J\| \quad \forall \quad x_I \neq x_J\} \quad (2.43)$$

A célula de Voronoi T_I é a região onde todos os pontos nela contidos, são mais próximos de x_I do que qualquer outro nó. Um diagrama de Voronoi de um conjunto de nós é a união de todas as células de Voronoi associadas a cada nó pertencente ao domínio. A Figura 2.8 (a) ilus-

tra um diagrama de Voronoi. Considerando o conjunto de nós, M , o menor polígono convexo que contém estes nós é chamado de feixe convexo, $CH(M)$ (*convex Hull*). As células associadas aos nós internos de $CH(M)$ são fechadas e convexas, e as células associadas aos nós da fronteira de $CH(M)$ são abertas. As arestas que delimitam as células de Voronoi são bissetores perpendiculares.

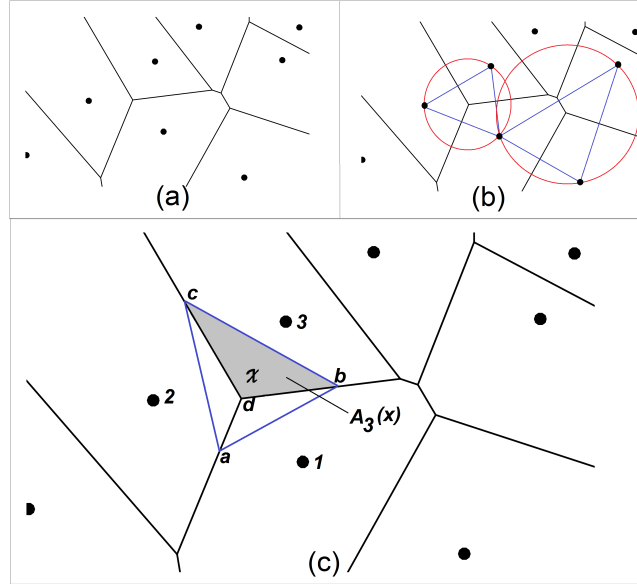


Figura 2.8: (a): Diagrama de Voronoi de primeira ordem de um conjunto de M nós. (b): Triangulação dual de Delaunay e circuncírculos. (c): Célula de Voronoi de segunda ordem de x .

2.2.4.3 Triangulação Dual de Delaunay

Como pode ser visto na Figura 2.8 (b), os triângulos de Delaunay são construídos pela conexão dos nós cujas células de Voronoi possuem fronteiras comuns. A triangulação de Delaunay é feita de forma que o círculo escrito ao triângulo de Delaunay, chamado de circuncírculo, não pode possuir nenhum nó em seu interior. No chamado caso não-degenerado, só existem três nós no bordo do círculo, os nós que formam os vértice do triângulo de Delaunay. O centro do circuncírculo é um vértice de Voronoi. Pelo critério do circuncírculo vazio: se a distância euclidiana entre um ponto x e o vértice de Voronoi é menor que o raio do circuncírculo, os vértices do triângulo de Delaunay são vizinhos naturais de x (Sukumar (1998a)).

2.2.4.4 Célula de Voronoi de Segunda Ordem

Seja Ω uma região no espaço Euclidiano, onde está distribuído um conjunto de M nós. A célula de Voronoi T_{IJ} associada aos nós x_I e x_J em Ω é dada por (Sukumar *et al.* (2001)):

$$T_{IJ} = \{x \in \mathbb{R} : \|x - x_I\| < \|x - x_J\| < \|x - x_K\| \quad \forall \quad x_K \neq x_I, x_J\} \quad (2.44)$$

A célula de Voronoi de segunda ordem, T_{IJ} , é a região onde todos os pontos nela contidos, são mais próximos de x_I e x_J do que qualquer outro nó, mas ainda são mais próximos de x_I . Para obter a célula de segunda ordem é inserido um ponto x no interior de um diagrama de Voronoi, depois é feita a tesselação de Voronoi como para os demais nós, como pode ser visto na Figura 2.8 (c). Nessa figura ainda são apresentados:

- O polígono abc é a célula de Voronoi de primeira ordem associada ao ponto $x(T_x)$.
- O polígono abd é a célula de Voronoi de segunda ordem associada ao ponto x e ao nó $x_1(T_{x_1})$.
- O polígono acd é a célula de Voronoi de segunda ordem associada ao ponto x e ao nó $x_2(T_{x_2})$.
- O polígono bcd é a célula de Voronoi de segunda ordem associada ao ponto x e ao nó $x_3(T_{x_3})$.

Os nós que compartilham as arestas da célula de Voronoi de x são chamados de vizinhos naturais de x .

2.2.4.5 Construção das Funções de Forma

Como citado anteriormente, as funções de forma do NEM são construídas com base no diagrama de Voronoi e na triangulação dual de Delaunay. Estas funções são definidas utilizando medidas de Lebesgue normalizadas (comprimento, área ou volume) (Sukumar *et al.* (2001); Sukumar (1998a)). Então, sendo $A(x)$ a área da célula de Voronoi de primeira ordem de x e $A_I(x)$ a área da célula de segunda ordem associada a x e ao nó x_I , a função de forma e suas derivadas parciais são dadas por:

$$\phi_I(x) = \frac{A_I(x)}{A(x)} \quad (2.45)$$

$$\phi_{I,j} = \frac{A_{I,j}(x) - \phi_I(x)A_{,j}(x)}{A(x)} \quad (2.46)$$

onde a notação $', j'$ define uma derivada parcial de primeira ordem na direção j . Percebe-se pelas equações (2.45) e (2.46) que as funções de forma e suas derivadas parciais dependem do diagrama de Voronoi, que por sua vez dependerá da distribuição dos nós no domínio e da

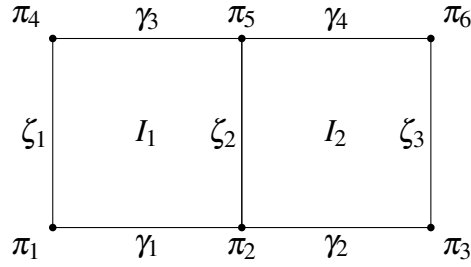


Figura 2.9: Representação topológica da distribuição dos nós no domínio com duas células adjacentes.

posição de x . Dentre os algoritmos conhecidos para o cálculo de ϕ e suas derivadas parciais está o algoritmo de [Watson \(1981\)](#), que foi implementado neste trabalho.

2.2.5 Imposição das Condições Periódicas

As condições naturais aparecem na forma fraca, mas as condições essenciais devem ser impostas. Uma forma seria construir funções de forma que sejam naturalmente periódicas, como em [Nicomedes *et al.* \(2012\)](#). Contudo, esse não é o caso das funções do FEM, S-FEM e do NEM. Porém, pelo fato de serem interpolantes, essa periodicidade pode ser facilmente imposta a todos os três os métodos. Para a função de teste v , basta que a triangulação que representa o domínio de integração seja periódica. Neste trabalho foram usadas duas células unitárias adjacentes como domínio, que é ilustrado topologicamente na Figura 2.9. Isso garante a periodicidade e ainda mantém o baixo custo computacional. No caso da função de forma u , é preciso impor a periodicidade dos coeficientes d_i 's, para isso foi feita uma fatoração.

Primeiramente, considere que a matriz $\mathbf{d}$ dos autovetores, pode ser representada por:

$$\mathbf{d} = [\mathbf{d}_{\pi_1}, \mathbf{d}_{\pi_2}, \mathbf{d}_{\pi_3}, \mathbf{d}_{\pi_4}, \mathbf{d}_{\pi_5}, \mathbf{d}_{\pi_6}, \mathbf{d}_{\gamma_1}, \mathbf{d}_{\gamma_2}, \mathbf{d}_{\gamma_3}, \mathbf{d}_{\gamma_4}, \mathbf{d}_{\zeta_1}, \mathbf{d}_{\zeta_2}, \mathbf{d}_{\zeta_3}, \mathbf{d}_{I_1}, \mathbf{d}_{I_2}]^T \quad (2.47)$$

$\mathbf{d}_{\pi_\ell}$ sendo os autovetores associados aos nós π_ℓ , com $\ell = 1, \dots, 6$. Observa-se que:

- $\mathbf{d}_{\gamma_\ell} \Rightarrow$ autovetores associados aos nós que residem em γ_ℓ , com $\ell = 1, 2, 3, 4$.
- $\mathbf{d}_{\zeta_\ell} \Rightarrow$ autovetores associados aos nós que residem em ζ_ℓ , com $\ell = 1, 2, 3$.
- $\mathbf{d}_{I_\ell} \Rightarrow$ autovetores associados aos nós que residem em I_ℓ , com $\ell = 1, 2$.

Aplicando as condições periódicas é possível reescrever a matriz $\mathbf{d}$ como:

$$\mathbf{d} = \mathbf{P}\mathbf{d}_f \quad (2.48)$$

onde:

$$\mathbf{d}_f = \begin{bmatrix} \mathbf{d}_{\pi_1} \\ \mathbf{d}_{\gamma_1} \\ \mathbf{d}_{\zeta_1} \\ \mathbf{d}_{I_1} \end{bmatrix} \quad (2.49)$$

$$\mathbf{P} = \begin{bmatrix} V^{ones} & 0 & 0 & 0 \\ 0 & I_\gamma & 0 & 0 \\ 0 & I_\gamma & 0 & 0 \\ 0 & I_\gamma & 0 & 0 \\ 0 & I_\gamma & 0 & 0 \\ 0 & 0 & I_\zeta & 0 \\ 0 & 0 & I_\zeta & 0 \\ 0 & 0 & I_\zeta & 0 \\ 0 & 0 & 0 & I_I \\ 0 & 0 & 0 & I_I \end{bmatrix}. \quad (2.50)$$

Na equação (2.50), V^{ones} é um vetor de 1's de dimensão 6, I representa matrizes identidade e 0 matrizes de zeros. As dimensões dessas matrizes devem ser escolhidas com cuidado, afim de garantir a consistência da fatoração. Reescrevendo a equação (2.42), e substituindo a forma fatorada de $\mathbf{d}$:

$$\mathbf{A}\mathbf{P}\mathbf{d}_f = \lambda_f \mathbf{B}\mathbf{P}\mathbf{d}_f \quad (2.51)$$

para conservar as características Hermitianas do sistema, é efetuado o produto, dos dois lados da equação (2.51), pela matriz transposta conjugada de $\mathbf{P}$:

$$(\mathbf{P}^H \mathbf{A} \mathbf{P}) \mathbf{d}_f = \lambda_f (\mathbf{P}^H \mathbf{B} \mathbf{P}) \mathbf{d}_f \quad (2.52)$$

se $\mathbf{A}_f = \mathbf{P}^H \mathbf{A} \mathbf{P}$ e $\mathbf{B}_f = \mathbf{P}^H \mathbf{B} \mathbf{P}$ é possível reescrever a equação (2.52) como:

$$(\mathbf{A}_f - \lambda_f \mathbf{B}_f) \mathbf{d}_f = 0. \quad (2.53)$$

É importante notar que essa fatoração reduz o número de incógnitas a uma quantidade menor que o número de graus de liberdade de uma única célula unitária. Portanto, dessa forma é possível inserir as condições essenciais, reduzir o número de incógnitas do sistema e ainda conservar o sistema Hermitiano. Uma observação deve ser feita acerca do uso desse procedimento: deve-se garantir que a triangulação seja periódica em a , garantindo que o número de nós nas fronteiras paralelas sejam iguais.

2.2.6 Validação Computacional

Além da curva ilustrada na Figura 2.3 que foi calculada com o FEM, nesta seção serão apresentados resultados encontrados com o NEM. Foram realizadas simulações com códigos implementados no Matlab, utilizando dois tipos de cristais bidimensionais com rede de simetria quadrada. O primeiro é formado por colunas dielétricas infinitas de raio $r = 0,2a$ e imersas em ar. O segundo cristal é formado por veias dielétricas com espessura $l = 0,164a$, também imersas em ar. Em ambos os cristais, o dielétrico usado tem constante dielétrica $\epsilon_r = 8.9$. Na Figura 2.10 são mostradas as estruturas bandas dos modos TE e TM para os dois cristais.

Os resultados foram comparados com soluções de referência encontradas na literatura. Na figura as curvas tracejadas em vermelho representam as bandas do modo TE, enquanto que as curvas contínuas em azul representam as bandas do modo TM. Os marcadores circulares e quadrados representam pontos selecionados nas curvas apresentadas em Joannopoulos *et al.* (2011), pg. 67-68 e 72. Uma grande concordância nos resultados pode ser observada.

Nos diagramas das figuras 2.10 (a) e (b), nota-se a ocorrência de *gaps* fotônicos com razão *gap-midgap* de 38% e 18.9% respectivamente. Em 2.10 (a) o *gap* ocorre no modo TM enquanto que em 2.10 (b) ocorre no modo TE, mas em ambos os casos ocorrem entre as duas primeiras bandas. Isto se deve principalmente a dois fatores: a ortogonalidade entre os autovetores e a tendência do campo elétrico de se concentrar nas regiões de maior constante dielétrica. Para satisfazer o teorema variacional, os autovetores devem minimizar o funcional energético. Por esse teorema, quando o campo elétrico se concentra em regiões de mais alta constante dielétrica, obtém-se modos de mais baixa frequência, minimizando o funcional energético.

Tomando como exemplo o cristal representado na Figura 2.10 (a), que é formado de colunas dielétricas ao longo da coordenada z . No modo TE o campo elétrico tem suas linhas paralelas ao plano xy dificultando sua concentração nas regiões de elevado ϵ_r . Já no modo TM, o campo elétrico tem suas linhas orientadas ao longo de z propiciando a concentração do campo elétrico nas colunas. Assim, na primeira banda do modo TM, a tendência é que o campo elétrico concentre a maior parte de sua energia nas regiões de elevado ϵ_r , mantendo suas frequências mais baixas. Contudo, pelo fato do perfil de campo da segunda banda ser ortogonal ao da primeira, uma grande parte da energia que era concentrada nas colunas, agora se concentra

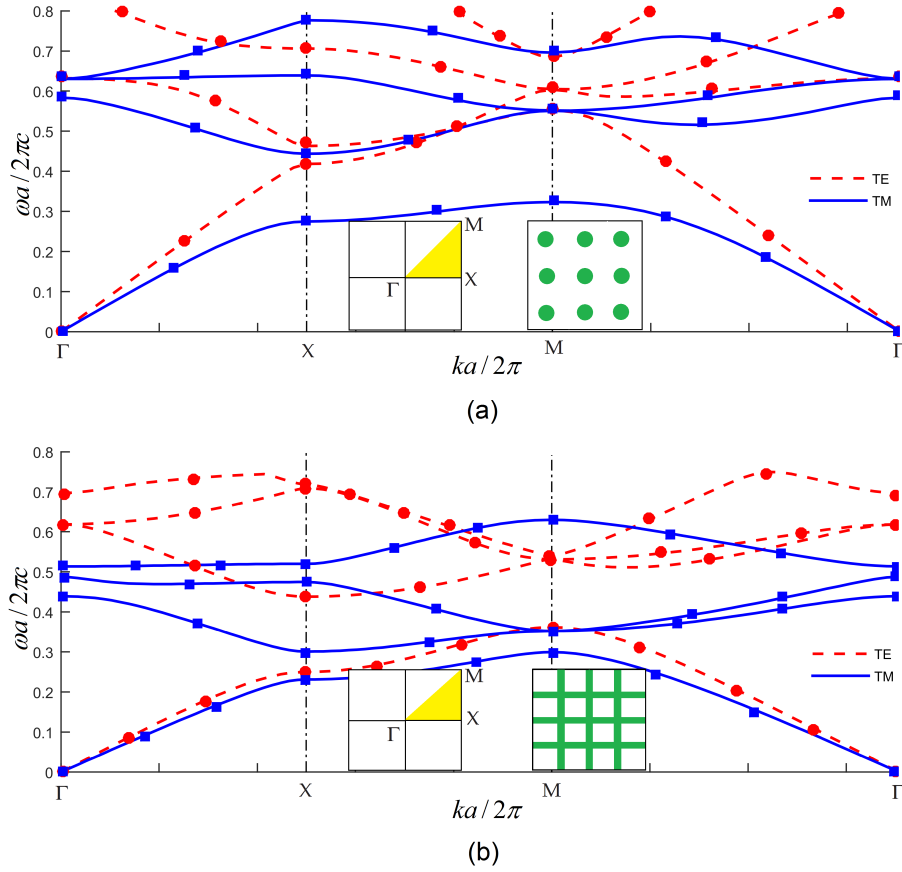


Figura 2.10: Estruturas de Bandas. (a): Estrutura de bandas de um cristal formado de colunas dielétricas com constante $\epsilon_r = 8,9$ e raio $r = 0,2a$. (b): Estrutura de bandas de um cristal formado de veias dielétricas com constante $\epsilon_r = 8,9$ e espessura $l = 0,165a$.

na região com ar. Isso faz com que as frequências associadas a banda 2 tenham um grande aumento, em relação a banda 1, causando o *gap* fotônico.

A Figura 2.11 traz a distribuição de energia de campo elétrico $\epsilon_r \mathbf{E}^2$ das bandas 1 e 2 ao longo de um domínio com nove células, para $k = \Gamma$, X e M. Observa-se que na banda 1, a energia está quase toda concentrada na região com constante dielétrica mais elevada. Isto faz com que suas frequências sejam baixas. Já na banda 2, observa-se que parte da energia é distribuída pela região de baixa constante dielétrica, fazendo com que suas frequências sejam bem maiores e gerando o *gap* fotônico. Os resultados estão de acordo com o esperado teórico, validando a formulação e a implementação.

Com o diagrama de estrutura de bandas as características de um cristal fotônico ficam determinadas e é possível projetar dispositivos com base nesse cristal. Contudo, para entender as características da propagação eletromagnéticas em um dispositivo baseado em cristal fotônico é necessário trabalhar no domínio do tempo.

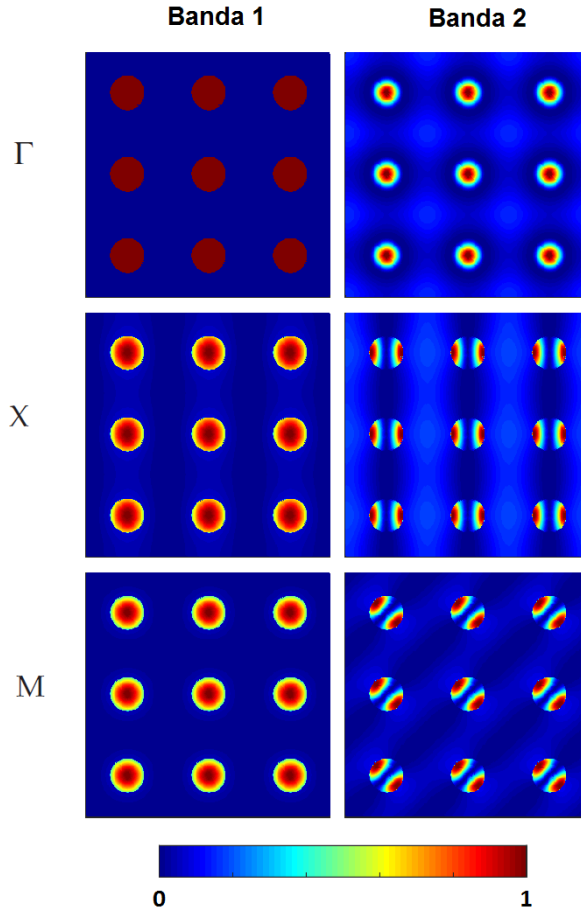


Figura 2.11: Energia: Distribuição da energia de campo elétrico $\epsilon_r \mathbf{E}^2(x,y)$ de um domínio formado por 9 células unitárias.

2.3 Simulações no Domínio do Tempo

Os métodos numéricos para solução de problemas eletromagnéticos mais gerais são as simulações no domínio do tempo. Estes métodos simulam a dependência temporal e espacial das equações de Maxwell. Tais métodos são ideais para problemas que variam com o tempo.

Métodos no domínio do tempo também podem ser usados na solução de problemas de autovalor. Para isso é utilizada a resposta em frequência de um pulso curto. Os autovalores podem ser identificados como picos no espectro de resposta. Este método também pode ser usado para identificar modos ressonantes ou *leaky*, onde a taxa de perda está associada a largura do pico. Na prática não é simplesmente olhar os picos na transformada de Fourier, técnicas avançadas de processamento de sinais devem ser aplicadas (Mandelshtam e Taylor (1997)). As vantagens desta técnica é que muitos autovalores podem ser computados de uma só vez e também é possível procurar por frequências em uma específica porção do espectro, como dentro do *gap* para modos associados a defeitos (Joannopoulos *et al.* (2011)).

Esses métodos também podem ser usados para resolver problemas no domínio da frequên-

cia, apresentando vantagens e desvantagens. A técnica mais comum de simulação no domínio do tempo é diferenças finitas no domínio do tempo. O FDTD divide o tempo e espaço em um *grid* de pontos discretos e aproxima as derivadas ($\nabla \times$ e $\partial/\partial t$) das equações de Maxwell por diferenças finitas. A integração do tempo é realizada geralmente usando o esquema "leap-frog". Nesse esquema $\mathbf{E}(t)$ é computado a partir de $\mathbf{E}(t - \Delta t)$ e $\mathbf{H}(t - \Delta t/2)$ será computado a partir de $\mathbf{H}(t + \Delta t/2)$ (Taflou e Hagness (2005)).

O método FDTD é frequentemente usado para calcular espectros de transmissão e reflexão, como os métodos de resposta em frequência. Contudo, os métodos no domínio do tempo são capazes de calcular a resposta de um sistema linear de muitas frequências de uma só vez. Isso é feito tomando a transformada de Fourier da resposta no tempo de um pulso curto. Porém, quando o interesse está em uma faixa estreita do espectro de frequência, o uso de um método no domínio do tempo não é adequado. Será necessário um tempo muito grande para computar um estreito espectro de frequência. Além disso, se o interesse está na resposta em estado estacionário devido a uma fonte harmônica $\mathbf{J}(\mathbf{x})e^{-j\omega t}$, ao usar um método no domínio do tempo é necessário esperar um grande tempo para que os efeitos do transiente desapareçam. Nesta situação um método de resposta em frequência promete ser mais eficiente (Joannopoulos *et al.* (2011)).

Um aspecto importante dos métodos de integração do tempo é o passo de tempo. Assim, se uma alta resolução espacial é requerida, uma alta resolução temporal é necessária, afim de manter a estabilidade numérica. O método de elementos finitos também pode ser usado para simulações no domínio do tempo. Para isso, requer um passo de tempo implícito. Isso significa que um sistema de equações $N \times N$ deve ser resolvido a cada passo de tempo, o que torna essa simulação muito cara do ponto de vista computacional. Como será visto no próximo capítulo, uma alternativa é o uso do DGTD, que traz as vantagens da discretização do FEM e apresenta custo computacional comparável ao FDTD.

Galerkin Descontínuo no Domínio do Tempo

Este capítulo trata da modelagem da propagação eletromagnética com o método DGTD, em meios dielétricos não dispersivos. Como o próprio nome sugere, este método é baseado em uma formulação local, onde os elementos são descontínuos. A continuidade da solução deve ser imposta pelo fluxo numérico. As condições de fronteira também são impostas através do fluxo numérico. O método DG foi primeiramente proposto na década de 70 por [Reed e Hill \(1973\)](#) para a aplicação na equação linear de transporte de nêutrons. O primeiro trabalho que trouxe uma formalização matemática para este método foi [Lesaint e Raviart \(1974\)](#), que também foi aplicado à equação linear de transporte de nêutrons. Ao longo dos anos, várias versões deste método foram desenvolvidas. Uma versão eficiente é descrita em [Hesthaven e Warburton \(2002b\)](#).

3.1 Equações de Maxwell

Considerando as equações rotacionais de Maxwell no sistema MKS (*Meter, Kilogram and Second*):

$$\begin{aligned}\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{H} &= \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}\end{aligned}\tag{3.1}$$

onde $\mathbf{H}$ e $\mathbf{E}$ são os campos magnético e elétrico, respectivamente. $\mathbf{D}$ é o campo de deslocamento elétrico e $\mathbf{B}$ é o campo de indução magnética. As equações constitutivas que relacionam $\mathbf{D}$ com $\mathbf{E}$, e $\mathbf{B}$ com $\mathbf{H}$, são:

$$\begin{aligned}\mathbf{D} &= \epsilon_0 \epsilon_r \mathbf{E} \\ \mathbf{B} &= \mu_0 \mu_r \mathbf{H}.\end{aligned}\tag{3.2}$$

Em (3.2), ϵ_0 e μ_0 são a permissividade elétrica e a permeabilidade magnética no espaço livre.

ϵ_r e μ_r são a permissividade relativa e a permeabilidade relativa do meio. Como o meio é um dielétrico $\mu_r = 1$. Considerando a ausência de corrente, $\mathbf{J} = 0$. A substituição de (3.2) em (3.1) leva a:

$$\begin{aligned}\nabla \times \mathbf{E} &= -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \\ \nabla \times \mathbf{H} &= \epsilon \frac{\partial \mathbf{E}}{\partial t}\end{aligned}\tag{3.3}$$

onde $\epsilon = \epsilon_0 \epsilon_r$.

As equações (3.3) estão em um formato geral tridimensional. Quando são considerados sistemas bidimensionais, por exemplo, em uma propagação de um modo TM, as equações podem ser reduzidas a:

$$\begin{aligned}\epsilon \frac{\partial E_z}{\partial t} &= \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \\ \mu_0 \frac{\partial H_x}{\partial t} &= -\frac{\partial E_z}{\partial y} \\ \mu_0 \frac{\partial H_y}{\partial t} &= \frac{\partial E_z}{\partial x}.\end{aligned}\tag{3.4}$$

As equações para o modo TE são análogas. Para um sistema unidimensional, as equações podem ser reduzidas a:

$$\begin{aligned}\epsilon \frac{\partial E_y}{\partial t} &= -\frac{\partial H_z}{\partial x} \\ \mu_0 \frac{\partial H_z}{\partial t} &= -\frac{\partial E_y}{\partial x}\end{aligned}\tag{3.5}$$

3.2 Lei de Conservação

Uma forma interessante de atacar as equações de Maxwell na forma hiperbólica é reescrevendo-as em forma de lei de conservação:

$$Q \partial_t \mathbf{q} + \nabla \cdot \mathbf{F} = 0\tag{3.6}$$

onde ∂_t representa uma diferenciação no tempo. De maneira geral, $\mathbf{q}$ é dado por:

$$\mathbf{q} = \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}. \quad (3.7)$$

O fluxo pode ser escrito da forma: $\mathbf{F} = (\mathbf{F}_x, \mathbf{F}_y, \mathbf{F}_z)^T$, sendo que,

$$\mathbf{F}_i = \begin{pmatrix} \hat{e}_i \times -\mathbf{H} \\ \hat{e}_i \times \mathbf{E} \end{pmatrix}. \quad (3.8)$$

O vetor unitário $\hat{e}_i$ define a direção de cada componente de campo de forma que $i = x, y, z$. A matriz Q fornece a informação sobre o meio:

$$Q = \begin{pmatrix} -\mu & 0 \\ 0 & \varepsilon \end{pmatrix}. \quad (3.9)$$

Trabalhar com as equações hiperbólicas, no formato de lei de conservação, proporciona o uso de esquemas TVD (*Total Variation Diminishing*, ou Variação Total Decrescente), como o proposto por [Harten \(1983\)](#), que possibilita um melhor controle de oscilações espúrias em problemas que envolvem rigidez (*stiffness*). Além disso, por ser um sistema conservativo, é possível obter um esquema simplético, mesmo depois da discretização. Para isso, é necessário a escolha de um fluxo numérico que conserve a energia. Como será visto mais a frente nesse texto, isso pode ser realizado com a escolha do chamado fluxo central.

3.2.1 Modo TM

Ao escrever o conjunto de equações em (3.4) como uma lei de conservação, $\mathbf{q} = [E_z, H_x, H_y]^T$, sendo:

$$Q = \begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & \mu_0 & 0 \\ 0 & 0 & \mu_0 \end{pmatrix}; \quad \nabla \cdot \mathbf{F}(\mathbf{q}) = \begin{pmatrix} -\frac{\partial H_y}{\partial x} + \frac{\partial H_x}{\partial y} \\ \frac{\partial E_z}{\partial y} \\ -\frac{\partial E_z}{\partial x} \end{pmatrix}.$$

O termo de fluxo pode ser reescrito como:

$$\begin{aligned}
\nabla \cdot \mathbf{F}(\mathbf{q}) &= -\partial_x \begin{pmatrix} H_y \\ 0 \\ E_z \end{pmatrix} + \partial_y \begin{pmatrix} H_x \\ E_z \\ 0 \end{pmatrix} \\
&= - \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \partial_x \mathbf{q} + \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \partial_y \mathbf{q} \\
&= A_x \partial_x \mathbf{q} + A_y \partial_y \mathbf{q}
\end{aligned} \tag{3.10}$$

3.2.2 Modo TEM

Agora, considerando o conjunto de equações unidimensionais dado por (3.5):

$$\mathbf{q} = \begin{pmatrix} E_y \\ H_z \end{pmatrix}; \quad Q = \begin{pmatrix} \varepsilon & 0 \\ 0 & \mu \end{pmatrix}; \quad \nabla \cdot \mathbf{F}(\mathbf{q}) = \partial_x \begin{pmatrix} H_z \\ E_y \end{pmatrix}$$

3.3 Forma Variacional

Para aplicar a modelagem DG, primeiramente o domínio computacional é discretizado a partir de uma malha não estruturada. Essa malha pode ser formada por tetraedros quando o problema discretizado é tridimensional. Caso o problema seja bidimensional, podem ser usados triângulos ou outros polígonos bidimensionais. Em problemas unidimensionais, o domínio é discretizado por segmentos de reta. Na discretização, os elementos não devem se sobrepor. Dessa forma, o domínio computacional é definido por: $\Omega = \bigcup_{k=1}^K \Omega_k$, sendo K o número de elementos, Ω_k , que formam a malha.

Na discretização espacial DG, os elementos devem ser descontínuos. Essa descontinuidade se reflete na forma variacional, que deve ser local. A fim de chegar na forma variacional, é definido um resíduo que é ponderado por uma função de teste $L_j(\mathbf{r}, t)$, definida localmente em cada elemento. Pelo método dos resíduos ponderados, a integral do resíduo em cada elemento deve ser nula:

$$\int_{\Omega_k} (Q \partial_t \mathbf{q}(\mathbf{r}, t) + \nabla \cdot \mathbf{F}(\mathbf{q})) L_j(\mathbf{r}, t) d\Omega = 0. \tag{3.11}$$

Aplicando o teorema da divergência a (3.11):

$$\int_{\Omega_k} (Q\partial_t \mathbf{q}(\mathbf{r}, t) L_j(\mathbf{r}, t) + \mathbf{F}(\mathbf{q}) \cdot \nabla L(\mathbf{r}, t)) d\Omega = - \int_{\Gamma_k} \hat{\mathbf{n}} \cdot \mathbf{F}(\mathbf{q}) L_j(\mathbf{r}, t) d\Gamma \quad (3.12)$$

que é a forma variacional fraca. É importante notar que (3.12) refere-se apenas a um elemento local Ω_k e não traz nenhuma informação sobre sua vizinhança. A fim de promover a continuidade da solução entre elementos vizinhos, o fluxo $\mathbf{F}$ em (3.12) é substituído por um fluxo numérico $\mathbf{F}^*$ e em seguida, o teorema da divergência é novamente aplicado. O resultado é a forma variacional forte:

$$\int_{\Omega_k} (Q\partial_t \mathbf{q}(\mathbf{r}, t) + \nabla \cdot \mathbf{F}(\mathbf{q})) L_j(\mathbf{r}, t) d\Omega = \int_{\Gamma_k} \hat{\mathbf{n}} \cdot (\mathbf{F}(\mathbf{q}) - \mathbf{F}^*(\mathbf{q}^-, \mathbf{q}^+)) L_j(\mathbf{r}, t) d\Gamma. \quad (3.13)$$

O fluxo numérico em (3.13) tem a função de promover a continuidade da solução, mas também pode ser usado para impor condições de fronteira e fontes. Portanto, o fluxo numérico em uma aresta deve ser função tanto do elemento local, $\mathbf{q}^-$, quanto do elemento vizinho, $\mathbf{q}^+$.

3.3.1 Fluxo Numérico

Existem algumas possibilidades para a escolha do fluxo numérico, contudo os mais usados são o fluxo central e o *upwind*, ou ainda, uma versão híbrida destes dois. A escolha do fluxo pode influenciar na precisão da solução. O uso do fluxo central traz um erro de dispersão maior, porém, o erro de dissipação é nulo. No caso do fluxo *upwind* o erro de dissipação existe, mas o erro de dispersão é significativamente menor que no fluxo central. Um fluxo estável e convergente é dado por [Hesthaven e Warburton \(2002b\)](#):

$$\hat{\mathbf{n}} \cdot (\mathbf{F}(\mathbf{q}) - \mathbf{F}^*(\mathbf{q}^-, \mathbf{q}^+)) = \begin{pmatrix} 1/\bar{Z}(\alpha[\Delta\mathbf{E} - \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \Delta\mathbf{E})] + Z^+ \hat{\mathbf{n}} \times \Delta\mathbf{H}) \\ 1/\bar{Y}(\alpha[\Delta\mathbf{H} - \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \Delta\mathbf{H})] - Y^+ \hat{\mathbf{n}} \times \Delta\mathbf{E}) \end{pmatrix} \quad (3.14)$$

onde $\Delta\mathbf{E} = \mathbf{E}^+ - \mathbf{E}^-$ e $\Delta\mathbf{H} = \mathbf{H}^+ - \mathbf{H}^-$. Na expressão do fluxo estão presentes as impedâncias: $Z^\pm = \sqrt{\mu^\pm/\epsilon^\pm}$ e $\bar{Z} = Z^- + Z^+$ e também as admitâncias: $Y^\pm = 1/Z^\pm$ e $\bar{Y} = Y^- + Y^+$. Caso $\alpha = 1$, o resultado é o fluxo *upwind*, $\alpha = 0$ resulta no fluxo central. Quando α assume um valor entre $[0, 1]$, resulta em um fluxo híbrido. No caso particular do modo TM, (3.14) pode ser

simplificada como:

$$\hat{\mathbf{n}} \cdot (\mathbf{F}(\mathbf{q}) - \mathbf{F}^*(\mathbf{q}^-, \mathbf{q}^+)) = \begin{pmatrix} 1/\bar{Z}(\alpha\Delta E_z + Z^+(n_x\Delta H_y - n_y\Delta H_x)) \\ 1/\bar{Y}(\alpha[\Delta H_x - n_x(n_x\Delta H_x + n_y\Delta H_y)] - Y^+n_y\Delta E_z) \\ 1/\bar{Y}(\alpha[\Delta H_z - n_y(n_x\Delta H_x + n_y\Delta H_y)] + Y^+n_x\Delta E_z) \end{pmatrix} \quad (3.15)$$

3.3.1.1 Condições de Fronteira

Uma parte importante da modelagem passa pela imposição das condições de fronteira, que podem ser pontos, linhas, ou superfícies e são caracterizadas dependendo do material que as compõem. Três tipos muito comuns são PEC (*perfect electrical conductor*), PMC (*perfect magnetic conductor*) e absorventes.

Em uma superfície PEC a componente tangencial do campo elétrico é nula, ou seja: $\hat{\mathbf{n}} \times \mathbf{E} = 0$. A componente tangencial do campo magnético é contínua, desde que não haja corrente elétrica na superfície. Assim, $\hat{\mathbf{n}} \times (\mathbf{H}_1 - \mathbf{H}_2) = \mathbf{J}_s = 0$. Para inserir as condições de fronteira no DG, supomos que exista um elemento externo à malha. Neste texto, este elemento será chamado de fantasma. Os valores dos campos do elemento fantasma, que residem na fronteira com o elemento real, podem ser escolhidos para que a interface entre estes elementos, represente um material PEC. É possível garantir que a componente tangencial do campo elétrico seja nula na fronteira, usando um princípio de espelhamento (Hesthaven e Warburton (2002b)). Para isso, basta impor que: $\mathbf{E}^+ = -\mathbf{E}^-$. Como $\Delta\mathbf{E} = \mathbf{E}^+ - \mathbf{E}^-$, em uma fronteira PEC, $\Delta\mathbf{E} = -2\mathbf{E}^-$. Para impor a continuidade da componente tangencial do campo magnético, basta fazer com que $\mathbf{H}^+ = \mathbf{H}^-$ e consequentemente $\Delta\mathbf{H} = 0$.

A condição de condutor magnético perfeito (PMC) é aplicada a superfícies onde se espera que o campo magnético seja totalmente normal a elas. Nesse caso, é necessário forçar a componente tangencial do campo magnético se anular nestas superfícies. Assim, para uma fronteira PMC, basta impor $\mathbf{H}^+ = -\mathbf{H}^-$. Isso implica em $\Delta\mathbf{H} = -2\mathbf{H}^-$. Já a componente tangencial do campo elétrico deve ser contínua na fronteira PMC. Essa continuidade é forçada impondo que $\mathbf{E}^+ = \mathbf{E}^-$, que tem como consequência: $\Delta\mathbf{E} = 0$.

A condição absorvente simula fronteiras livres. Quando a onda incide nesta fronteira, é absorvida. Existem algumas formas de impor condições absorventes. Por exemplo, o uso de uma PML (*perfectly matched layer*), que é o estado da arte em condições absorventes. Uma outra forma de impor condição absorvente, é através da condição de Silver Müller. Esta condição é bastante usada, principalmente no método dos volumes finitos, onde é imposta localmente com fluxo numérico. De acordo com Rao (1999), a condição Silver-Müller leva à seguinte relação entre os campos:

$$\sqrt{\frac{\epsilon_0}{\mu_0}} \hat{\mathbf{n}} \times \mathbf{E} + \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{H}) = 0 \quad (3.16)$$

É importante notar que: para garantir a relação dada em (3.16), basta impor que as componentes tangenciais dos campos, $\hat{\mathbf{n}} \times \mathbf{E}$ e $\hat{\mathbf{n}} \times \mathbf{H}$, sejam nulas. Isso implica que no fluxo numérico: $\Delta \mathbf{E} = -2\mathbf{E}^-$ e $\Delta \mathbf{H} = -2\mathbf{H}^-$.

Dessa forma, no DG as condições de fronteira podem ser impostas de forma fraca através do fluxo numérico. Para tal, basta substituir as diferenças de campo, como na tabela 3.1 encontrada em [Busch et al. \(2011\)](#), no fluxo.

Tabela 3.1: Condições de fronteira

Tipo de fronteira	$\Delta \mathbf{E}$	$\Delta \mathbf{H}$
PEC	$-2\mathbf{E}^-$	0
PMC	0	$-2\mathbf{H}^-$
Silver Müller (primeira ordem)	$-2\mathbf{E}^-$	$-2\mathbf{H}^-$

3.3.1.2 Fontes

Muitas vezes é necessário excitar um campo em alguma parte do domínio computacional, como é o caso de um guia de ondas. Nas simulações, pretende-se excitar a entrada de um guia com um campo eletromagnético. Isso pode ser feito usando o fluxo numérico, de forma análoga a usada para impor as condições PEC, PMC e absorvente. Para isso, basta garantir que os valores de campo dos nós do elemento fantasma, nesta fronteira, sejam os valores dos campos a serem excitados. Em outras palavras, para excitar a entrada do guia com um campo elétrico $\mathbf{E}(t)$ e um campo magnético $\mathbf{H}(t)$, basta impor: $\mathbf{E}^+ = \mathbf{E}(t)$ e $\mathbf{H}^+ = \mathbf{H}(t)$. Dessa forma, é possível dividir o vetor fluxo em uma soma de dois vetores. O primeiro, garante a conexão entre os elementos internos e impõe as condições de fronteira, e o segundo, insere a fonte do sinal: $F^{fluxo\ total} = F^{fluxo} + S^{fonte}$. Uma outra forma de adicionar fontes é através da técnica TF/SF (*total field/scattered field*), uma descrição dessa técnica aplicada ao DG pode ser encontrada em [Busch et al. \(2011\)](#). As fontes também podem ser impostas de forma forte por condições de Dirichlet.

Por vezes a excitação de estruturas pode não ser uma tarefa simples. Cada forma apresenta vantagens e desvantagens. A imposição através do fluxo numérico pode trazer alguma imprecisão. Por exemplo se a intenção é usar duas ou mais fontes coerentes, podem ocorrer diferença de amplitude entre as ondas geradas. A imposição da fonte de forma forte exclui esse problema, porém acrescenta outro. Quando alguma energia é refletida para a região da fonte a solução nessa região fica imprecisa.

3.4 Forma Semi-discreta

A fim de chegar à forma semi-discreta, é aplicado o método de Galerkin. Nesse caminho, supomos que a solução local pode ser representada por:

$$\mathbf{q}_k^h = \sum_{i=1}^{N_p} q_i(\mathbf{r}_i, t) L_i(\mathbf{r}) = \sum_{n=1}^{N_p} \hat{q}_n \psi_n(\mathbf{r}), \quad (3.17)$$

onde as funções L_i formam uma base nodal e ψ_n formam uma base modal. Os coeficientes nodais e modais são relacionados da seguinte forma (Hesthaven e Warburton (2002b)):

$$\mathbf{q}_{nodais} = \mathcal{V} \hat{\mathbf{q}}_{modais}, \quad (3.18)$$

sendo $\mathcal{V}$ a matriz generalizada de Vandermonde, $\mathcal{V}_{ij} = \psi_j(\mathbf{r}_i)$. Uma opção muito comum para as funções L_i são os polinômios de Lagrange. Uma grande vantagem em usar essas funções é a propriedade da delta de Kronecker:

$$L_i(\mathbf{r}_j) = \begin{cases} 1, & \text{se } i = j, \\ 0, & \text{se } i \neq j. \end{cases} \quad (3.19)$$

Isso implica que essas funções são interpolantes e condições essenciais podem ser impostas diretamente. Além disso, é possível construir bases espectrais utilizando polinômios de Lagrange. Para isso basta que os pontos de Lagrange coincidam com os pontos usados na quadratura.

Polinômios de Lagrange são formados pela combinação linear de monômios e podem ser representados de forma geral como (Busch *et al.* (2011)):

$$L_i(\mathbf{r}) = \sum_{k,l,m=0}^{k+l+m \leq p} a_{k,l,m}^i x^k y^l z^m, \quad (3.20)$$

onde p representa o grau do polinômio. O número de nós devem respeitar a seguinte relação:

$$Np = p + 1 \quad (1D)$$

$$Np = \frac{1}{2}(p+1)(p+2) \quad (2D)$$

$$Np = \frac{1}{6}(p+1)(p+2)(p+3) \quad (3D)$$

As funções modais, ψ_n , devem ser ortogonais. Quando um polinômio, P_n , é ortogonal a outro polinômio, P_m (Chihara (2011)):

$$\int_{\Omega} P_n(\mathbf{r}) P_m(\mathbf{r}) w(\mathbf{r}) d\Omega = \begin{cases} Cte, & \text{se } n = m \\ 0, & \text{caso contrário,} \end{cases} \quad (3.21)$$

onde w é uma função peso e n e m representam o grau do polinômio P . Quando $Cte = 1$, P é dito ortonormal. Na literatura são encontrados vários exemplos de polinômios ortogonais, como Chebyshev, Jacobi e Laguerre. A base usada nesse trabalho foi desenvolvida por Hesthaven e Warburton (2002b). Considerando um espaço definido pelas coordenadas (r, s) , uma base canônica pode ser definida como:

$$\begin{aligned} \psi_m(\mathbf{r}) &= r^i s^j, \quad (i, j) \geq 0, (i + j) \leq N \\ m &= j + (N + 1)i + 1 - \frac{1}{2}(i - 1). \end{aligned} \quad (3.22)$$

A aplicação do processo de ortonormalização de Gram Schmidt leva a uma base completa de polinômios ortonormais:

$$\psi_m(\mathbf{r}) = \sqrt{2} P_i(a) P_j^{(2i+1, 0)}(b) (1 - b)^i, \quad (3.23)$$

sendo

$$a = 2 \frac{1 + r}{1 - s} - 1; \quad b = s, \quad (3.24)$$

e $P_n^{(\alpha, \beta)}$ são os polinômios de Jacobi. Quando $\alpha = \beta = 0$, (3.23) são os polinômios de Legendre.

A distribuição dos N_p nós é feita em um elemento de referência e depois são mapeados para o elemento físico. A distribuição nodal deve ser feita de forma que garanta o bom condicionamento das matrizes que resultarão da discretização espacial. Por isso, uma distribuição equidistante deve ser evitada. Na literatura são encontradas várias formas de se obter uma boa distribuição nodal (Chen e Babuška (1995); Hesthaven (1998); Taylor *et al.* (2000)). Apesar desses esquemas levarem a boas distribuições nodais, normalmente eles são custosos computacionalmente. Geralmente, as distribuições nodais para uma gama de ordens polinomiais é obtida previamente e tabelada. Em Hesthaven e Warburton (2002b) foi desenvolvida uma forma relativamente simples de obter boas distribuições nodais. Nesse processo uma distribuição equidistante de nodos em um simplex é mapeada para uma distribuição Legendre-Gauss-Lobato.

Esse processo é rápido e pode eliminar o uso de tabelas. Ou seja, a distribuição é feita para uma dada ordem polinomial no momento da simulação, sem gerar um aumento significativo no custo computacional. Isso facilita o uso de ordens polinomiais muito altas.

3.4.1 Modo TM

Aqui será desenvolvida a discretização no domínio do espaço para o modo TM. Primeiramente, considere o lado esquerdo de (3.6) que está relacionado somente ao elemento local. Lembrando que o fluxo pode ser escrito como em (3.10):

$$\int_{\Omega^k} (Q^k \partial_t \mathbf{q}^k + \nabla \cdot \mathbf{F}(\mathbf{q}^k)) L_j^k d\Omega = \int_{\Omega^k} (Q^k \partial_t \mathbf{q}^k + A_x \partial_x \mathbf{q}^k + A_y \partial_y \mathbf{q}^k) L_j^k d\Omega \quad (3.25)$$

Para cada variável no sistema de equações (3.10) é feita uma aproximação por funções de base, de forma que: $u_h^k = \sum_{i=1}^{N_p} u_i^{k,i} L_i^k(x, y)$, onde N_p é o número de incógnitas para cada campo. Os coeficientes são funções do tempo e a função de base, somente do espaço. Para que o processo de discretização seja melhor entendido, considere cada equação do conjunto dado por (3.10), separadamente. Utilizando a expansão por funções de base chega-se a:

$$\begin{aligned} \int_{\Omega^k} (\epsilon^k \partial_t E_z^k - \partial_x H_y^k + \partial_y H_x^k) L_j^k d\Omega &= \epsilon^k M^k \partial_t \mathbf{E}_z^k - [K^{k,x}]^T \mathbf{H}_y^k + [K^{k,y}]^T \mathbf{H}_x^k \\ \int_{\Omega^k} (\mu_0^k \partial_t H_x^k + \partial_y E_z^k) L_j^k d\Omega &= \mu_0^k M^k \partial_t \mathbf{H}_x^k + [K^{k,y}]^T \mathbf{E}_z^k \\ \int_{\Omega^k} (\mu_0^k \partial_t H_y^k - \partial_x E_z^k) L_j^k d\Omega &= \mu_0^k M^k \partial_t \mathbf{H}_y^k - [K^{k,x}]^T \mathbf{E}_z^k \end{aligned} \quad (3.26)$$

onde $\mathbf{E}_z^k$, $\mathbf{H}_x^k$ e $\mathbf{H}_y^k$ são vetores, cada um com dimensão $N_p \times 1$, que contém os valores nodais dos campos. A permissividade elétrica e a permeabilidade magnética são consideradas constantes no interior do elemento k . As matrizes M^k , $K^{k,x}$ e $K^{k,y}$, cada uma com dimensão $N_p \times N_p$, têm suas entradas definidas por:

$$\begin{aligned} M_{i,j}^k &= \int_{\Omega^k} L_i^k L_j^k d\Omega \\ K_{i,j}^{k,x} &= \int_{\Omega^k} \partial_x L_i^k L_j^k d\Omega \\ K_{i,j}^{k,y} &= \int_{\Omega^k} \partial_y L_i^k L_j^k d\Omega. \end{aligned} \quad (3.27)$$

Sendo as funções L_i polinômios de Lagrange, por causa da relação entre as base nodais e modais que foi mencionada anteriormente, a matriz de massa pode ser calculada da seguinte forma:

$$M^k = J^k (\mathcal{V} \mathcal{V}^T)^{-1}, \quad (3.28)$$

onde J^k é o jacobiano do elemento k . É importante observar que, como os polinômios ψ_n são ortonormais, a matriz de massa é diagonal. As matrizes de rigidez também podem ser calculadas a partir da base modal. Primeiramente considere que:

$$\frac{\partial}{\partial x} = \frac{\partial r}{\partial x} \mathcal{D}_r + \frac{\partial s}{\partial x} \mathcal{D}_s, \quad \frac{\partial}{\partial y} = \frac{\partial r}{\partial y} \mathcal{D}_r + \frac{\partial s}{\partial y} \mathcal{D}_s.$$

As coordenadas (r, s) definem o espaço de referência e as matrizes $\mathcal{D}_r$ e $\mathcal{D}_s$ produzem uma diferenciação em relação a r e s respectivamente. Reescrevendo as matrizes de rigidez:

$$\begin{aligned} K^{k,x} &= \left(\frac{\partial r}{\partial x} \mathcal{D}_r + \frac{\partial s}{\partial x} \mathcal{D}_s \right) (\mathcal{V} \mathcal{V}^T)^{-1} \\ K^{k,y} &= \left(\frac{\partial r}{\partial y} \mathcal{D}_r + \frac{\partial s}{\partial y} \mathcal{D}_s \right) (\mathcal{V} \mathcal{V}^T)^{-1}. \end{aligned} \quad (3.29)$$

Para determinar as matrizes $\mathcal{D}_r$ e $\mathcal{D}_s$, considere que:

$$\begin{aligned} \mathcal{V}_{r,(i,j)} &= \frac{\partial \psi_j(r_i, s_i)}{\partial r} \\ \mathcal{V}_{s,(i,j)} &= \frac{\partial \psi_j(r_i, s_i)}{\partial s}. \end{aligned} \quad (3.30)$$

Além disso, $\mathcal{V}_r = \mathcal{D}_r \mathcal{V}$ e $\mathcal{V}_s = \mathcal{D}_s \mathcal{V}$ e portanto:

$$\begin{aligned} \mathcal{D}_r &= \mathcal{V}_r \mathcal{V}^{-1} \\ \mathcal{D}_s &= \mathcal{V}_s \mathcal{V}^{-1}. \end{aligned} \quad (3.31)$$

Com o objetivo de sintetizar a notação, os vetores que contém os coeficientes de cada

campo podem ser arranjados em um único vetor:

$$\mathbf{q}_t^k = \begin{pmatrix} \mathbf{E}_z^k \\ \mathbf{H}_x^k \\ \mathbf{H}_y^k \end{pmatrix}_{3N \times 1} \quad (3.32)$$

que nos permite escrever (3.25) na forma matricial:

$$\int_{\Omega^k} (Q^k \partial_t \mathbf{q}^k + A_x \partial_x \mathbf{q}^k + A_y \partial_y \mathbf{q}^k) L_j^k d\Omega = Q_k \mathcal{M}_k \partial_t \mathbf{q}_t^k + (\mathcal{A}_x \mathcal{K}_{k,x}^T + \mathcal{A}_y \mathcal{K}_{k,y}^T) \mathbf{q}_t^k \quad (3.33)$$

onde as matrizes $\mathcal{M}_k$, $\mathcal{K}_{k,x}$ e $\mathcal{K}_{k,y}$ são bloco diagonais, cada uma com dimensão $3N \times 3N$. Cada bloco é formado pelas matrizes M^k , $K^{k,x}$ e $K^{k,y}$, respectivamente. As matrizes $\mathcal{A}_x$ e $\mathcal{A}_y$ são extensões de A_x e A_y , de forma que:

$$\mathcal{A}_x = - \begin{pmatrix} 0_{N \times N} & 0_{N \times N} & I_N \\ 0_{N \times N} & 0_{N \times N} & 0_{N \times N} \\ I_N & 0_{N \times N} & 0_{N \times N} \end{pmatrix}; \quad \mathcal{A}_y = \begin{pmatrix} 0_{N \times N} & I_N & 0_{N \times N} \\ I_N & 0_{N \times N} & 0_{N \times N} \\ 0_{N \times N} & 0_{N \times N} & 0_{N \times N} \end{pmatrix} \quad (3.34)$$

Analogamente, Q_k é uma extensão de Q^k , onde ε e μ_0 são substituídos por εI_N e $\mu_0 I_N$, sendo I_N matrizes identidades de dimensão N . Assim, o lado esquerdo de (3.6) fica discretizado em relação ao espaço, e isso permite escrever:

$$Q_k \mathcal{M}_k \partial_t \mathbf{q}_t^k + (\mathcal{A}_x \mathcal{K}_{k,x}^T + \mathcal{A}_y \mathcal{K}_{k,y}^T) \mathbf{q}_t^k = \text{fluxo} \quad (3.35)$$

agora, para encontrar a forma semi-discreta basta discretizar o termo de fluxo, do lado direito de (3.6). Primeiramente, considere que:

$$\hat{\mathbf{n}} \cdot (\mathbf{F}(\mathbf{q}) - \mathbf{F}^*(\mathbf{q}^-, \mathbf{q}^+)) = F \Delta \mathbf{q} \quad (3.36)$$

mudando a notação, $(-, +) \rightarrow (k, m)$, para representar o elemento local e seu vizinho:

$$F = \begin{pmatrix} \alpha/\bar{Z} & -n_y Z^m/\bar{Z} & n_x Z^m/\bar{Z} \\ -n_y Y^m/\bar{Y} & \alpha(1 - n_x^2)/\bar{Y} & \alpha(1 - n_x n_y)/\bar{Y} \\ n_x Y^m/\bar{Y} & \alpha(1 - n_x n_y)/\bar{Y} & \alpha(1 - n_y^2)/\bar{Y} \end{pmatrix} \quad (3.37)$$

substituindo (3.36) na integral de fluxo em (3.6):

$$\int_{\Gamma_{\Omega^k}} \hat{\mathbf{n}} \cdot (\mathbf{F}(\mathbf{q}^k) - \mathbf{F}^*(\mathbf{q}^-, \mathbf{q}^+)) L_j^k d\Gamma = \int_{\Gamma_{\Omega^k}} F \Delta \mathbf{q} L_j^k d\Gamma \quad (3.38)$$

Em um elemento triangular, as três arestas são as fronteiras. Cada aresta é uma interseção com um elemento m , ou seja, $\Gamma_e = \Omega^k \cap \Omega^{m_e}$ portanto:

$$\int_{\Gamma_{\Omega^k}} F \Delta \mathbf{q} L_j^k d\Gamma = \sum_{e=1}^3 \int_{\Gamma_e} F_e (\mathbf{q}^{m_e} - \mathbf{q}^k) L_j^k d\Gamma \quad (3.39)$$

Discretizando $\mathbf{q}^k$ e $\mathbf{q}^{m_e}$ como antes chega-se a:

$$\sum_{e=1}^3 \int_{\Gamma_e} F_e (\mathbf{q}^{m_e} - \mathbf{q}^k) L_j^k d\Gamma = \sum_{e=1}^3 F_e \mathcal{M}_{1d}^e (\mathbf{q}_t^{m_e} - \mathbf{q}_t^k) \quad (3.40)$$

onde $\mathcal{M}_{1d}^e$ é uma matriz bloco diagonal com dimensão $3N \times 3N$ com cada bloco sendo a matriz de massa M^e , com entradas dada por:

$$M_{i,j}^e = \int_{\Gamma_e} L_i^k L_j^k d\Gamma \quad (3.41)$$

É importante frisar que em (3.41) há uma integral de linha sobre uma aresta e . Também é possível calcular a matriz M^e a partir de uma base modal. O traço das funções ψ_n nas arestas do elemento também são polinômios ortonormais, ψ_m^e . As funções em uma aresta e são (Hesthaven e Warburton (2002b)):

$$\psi_m^e(\mathbf{r}) = \frac{P_{m-1}(\mathbf{r})}{\sqrt{\gamma_m - 1}}, \quad (3.42)$$

onde P_m são polinômios de Legendre e

$$\gamma_m = \frac{2}{2m+1} \quad (3.43)$$

é o fator de normalização. Considerando que a matriz generalizada de Vandermonde unidimensional é dada por: $\mathcal{V}_{1d,(i,j)} = \psi_j(\mathbf{r}_i)$, a matriz de massa em cada aresta pode ser calculada por:

$$M^e = \frac{h^e}{2} (\mathcal{V}_{1d} \mathcal{V}_{1d}^T)^{-1}, \quad (3.44)$$

onde h^e é o comprimento da aresta. A forma semi-discreta para o modo TM fica totalmente definida:

$$\mathcal{Q}_k \mathcal{M}_k \partial_t \mathbf{q}_t^k + (\mathcal{A}_x \mathcal{K}_{k,x}^T + \mathcal{A}_y \mathcal{K}_{k,y}^T) \mathbf{q}_t^k = \sum_{e=1}^3 F_e \mathcal{M}_{1d}^e (\mathbf{q}_t^{m_e} - \mathbf{q}_t^k) \quad (3.45)$$

3.4.2 Modo TEM

Agora considere um modo TEM_x , descrito pela equação (3.5). A forma semi-discreta para esse problema pode ser desenvolvida de forma análoga ao modo TE. Seguindo os mesmos passos chega-se a:

$$\begin{aligned} \varepsilon M_{1d}^k \partial_t \mathbf{E}_y^k + K_{k,x}^T \mathbf{H}_z^k &= \mathbf{F}_{E_y}^{k,m} \\ \mu M_{1d}^k \partial_t \mathbf{H}_z^k + K_{k,x}^T \mathbf{E}_y^k &= \mathbf{F}_{H_z}^{k,m}, \end{aligned} \quad (3.46)$$

onde $\mathbf{F}^{k,m}$ representa um vetor fluxo numérico, e

$$M_{1d}^k = \frac{h^k}{2} \int_{x_i^k}^{x_f^k} \ell_i^k \ell_j^k dx; \quad K_{k,x} = \int_{x_i^k}^{x_f^k} \partial_x \ell_i^k \ell_j^k dx. \quad (3.47)$$

As funções $\ell(x)$ são polinômios de Lagrange unidimensionais. Em (3.47), h_k é o comprimento do segmento de reta e x_i^k e x_f^k são os vértices. As matrizes podem ser calculadas usando a base modal, de forma que:

$$\begin{aligned} M_{1d}^k &= \frac{h^k}{2} (\mathcal{V}_{1d} \mathcal{V}_{1d}^T)^{-1}, \\ K_{k,x} &= \frac{dr}{dx} \mathcal{D}_{r,1d} (\mathcal{V}_{1d} \mathcal{V}_{1d}^T)^{-1}. \end{aligned} \quad (3.48)$$

Analogamente ao caso bidimensional, $\mathcal{D}_{r,1d} = \mathcal{V}_{1d}^{-1} \mathcal{V}_{r,1d}$, sendo que:

$$\mathcal{V}_{r,1d,(i,j)} = \frac{d\psi_{1d,j}(x_i)}{dx}. \quad (3.49)$$

Ainda é preciso determinar o fluxo numérico a ser usado para que a forma semi-discreta fique totalmente definida. Em problemas de propagação eletromagnética, os fluxos mais utilizados são o central e o *upwind*. Em primeiro lugar, considerando o fluxo central:

$$f^*(E^k, E^m) = \frac{H^k + H^m}{2}; \quad f^*(H^k, H^m) = \frac{E^k + E^m}{2}. \quad (3.50)$$

Portanto, com o fluxo central os vetores em (3.46) ficam assim:

$$(\mathbf{F}_{E_y}^{k,m})^T = \begin{pmatrix} H_{z,N}^m \\ -H_{z,1}^k \\ \vdots \\ 0 \\ \vdots \\ H_{z,N}^k \\ -H_{z,1}^m \end{pmatrix}; \quad (\mathbf{F}_{H_z}^{k,m})^T = \begin{pmatrix} E_{y,N}^m \\ -E_{y,1}^k \\ \vdots \\ 0 \\ \vdots \\ E_{y,N}^k \\ -E_{y,1}^m \end{pmatrix}. \quad (3.51)$$

Agora considere o fluxo *upwind* dado por:

$$\begin{aligned} f^*(E_y^k, E_y^m) &= \frac{\bar{Z}^{-1}}{2} (\{ZH_z\} + \llbracket E_y \rrbracket); \\ f^*(H_z^k, H_z^m) &= \frac{\bar{Y}^{-1}}{2} (\{YE_y\} + \llbracket H_z \rrbracket) \end{aligned} \quad (3.52)$$

onde $Z^k = \sqrt{\mu^k/\epsilon^k}$ e $Y^k = 1/Z^k$, são a impedância e a admitância no elemento k . $\bar{Z} = (Z^k + Z^m)/2$ e $\bar{Y} = (Y^k + Y^m)/2$, são as médias destas grandezas entre elementos vizinhos. No caso particular de um meio homogêneo, $Z^k = Z^m = \bar{Z}$ e $Y^k = Y^m = \bar{Y}$, e por isso, a impedância e a admitância serão tratadas apenas como Z e Y . Em (3.52), $\{q\} = q^k + q^m$ e $\llbracket q \rrbracket = q^k - q^m$. Então, com o uso do fluxo *upwind*:

$$[(f(E_y^k) - f^*(E_y^k, E_y^m))\ell_j^k]_{x_1^k}^{x_N^k} = \begin{cases} \frac{1}{2} \left((H_{z,N}^m - H_{z,1}^k) - Z^{-1} (E_{y,N}^m - E_{y,1}^k) \right); & j = 1, \\ 0; & 1 < j < N, \\ \frac{1}{2} \left((H_{z,N}^k - H_{z,1}^m) + Z^{-1} (E_{y,N}^k - E_{y,1}^m) \right); & j = N. \end{cases} \quad (3.53)$$

$$[(f(H_z^k) - f^*(H_z^k, H_z^m))\ell_j^k]_{x_1^k}^{x_N^k} = \begin{cases} \frac{1}{2} \left((E_{y,N}^m - E_{y,1}^k) - Y^{-1} (H_{z,N}^m - H_{z,1}^k) \right); & j = 1, \\ 0; & 1 < j < N, \\ \frac{1}{2} \left((E_{y,N}^k - E_{y,1}^m) + Y^{-1} (H_{z,N}^k - H_{z,1}^m) \right); & j = N. \end{cases} \quad (3.54)$$

Assim, a forma semi-discreta com o fluxo *upwind* também fica totalmente determinada. O conjunto de equações (3.46) podem ser agrupadas da forma:

$$\mathcal{Q}^k \mathcal{M}_k \partial_t \mathbf{q}_t^k + \mathcal{K}_k \mathcal{R} \mathbf{q}_t^k = \mathbf{F}^{k,m}, \quad (3.55)$$

sendo

$$\begin{aligned} \mathbf{q}_t^k &= \begin{pmatrix} \mathbf{E}_y^k \\ \mathbf{H}_z^k \end{pmatrix}; \quad \mathcal{Q}^k = \begin{pmatrix} \varepsilon \mathbf{I}_{N \times N} & 0_{N \times N} \\ 0_{N \times N} & \mu \mathbf{I}_{N \times N} \end{pmatrix}; \quad \mathbf{F}^{k,m} = \begin{pmatrix} \mathbf{F}_{E_y}^{k,m} \\ \mathbf{F}_{H_z}^{k,m} \end{pmatrix}; \\ \mathcal{M}_k &= \begin{pmatrix} \mathbf{M}_k & 0_{N \times N} \\ 0_{N \times N} & \mathbf{M}_k \end{pmatrix}; \quad \mathcal{K}_k = \begin{pmatrix} \mathbf{K}_{x,k}^T & 0_{N \times N} \\ 0_{N \times N} & \mathbf{K}_{x,k}^T \end{pmatrix}; \quad \mathcal{R} = \begin{pmatrix} 0_{N \times N} & \mathbf{I}_{N \times N} \\ \mathbf{I}_{N \times N} & 0_{N \times N} \end{pmatrix}. \end{aligned}$$

3.4.2.1 Testes e Resultados

Como mencionado anteriormente, os fluxos numéricos mais usados em problemas de propagação, são o central e o *upwind*. A fim de fazer uma comparação qualitativa desses dois esquemas, uma implementação no Matlab para simular a propagação eletromagnética do modo TEM_x no ar, foi desenvolvida. A entrada do meio é excitada com um campo eletromagnético, de forma que: $E_{y,N}^0 = \sin(\omega t) e^{(t-t_0)^2/2\sigma^2}$ e $H_{z,N}^0 = (1/Z)E_{y,N}^0$. A frequência central do sinal simulado é $f_0 = 0,59 \text{ PHz}$ e $\sigma = 1.8 \times 10^{-15}$. Uma fronteira PEC foi colocada no extremo direito do domínio. Esta condição foi imposta de forma fraca, no fluxo numérico. Para isso, um princípio de espelhamento pode ser usado: $E_{y,N}^{N_e} = -E_{y,1}^{N_e+1}$ e $H_{z,N}^{N_e} = H_{z,1}^{N_e+1}$. É esperado que os campos sejam refletidos pela fronteira PEC. O campo magnético deve sofrer uma inversão de fase ao ser refletido e o campo elétrico não deve sofrer alteração em sua fase. Isto é verificado na Figura 3.1, que foi gerada com um tempo de propagação, $t = 5 \times 10^{-14} \text{ s}$.

A fim de comparar os fluxos central e *upwind*, a propagação do pulso em uma janela de, $t = 8,33 \times 10^{-14} \text{ s}$ foi simulada. As simulações foram realizadas com ordens polinomiais, $p = 2$ e $p = 10$. O resultado é demonstrado na figura 3.2. Quando $p = 10$, a solução com ambos

os fluxos é praticamente igual. Contudo, quando $p = 2$, a solução com fluxo central sofre uma forte distorção, enquanto que a solução com fluxo *upwind* não apresenta tanta distorção. Por outro lado, a solução com *upwind* sofre uma maior dissipação. Apesar de ser simplético, o fluxo central causa uma alta dispersão numérica à propagação. O fluxo *upwind* é dissipativo, porém é muito menos dispersivo. É visto também na figura 3.2 que tanto a dispersão quanto a dissipação são fortemente amenizadas com o uso de altas ordens de aproximação espacial.

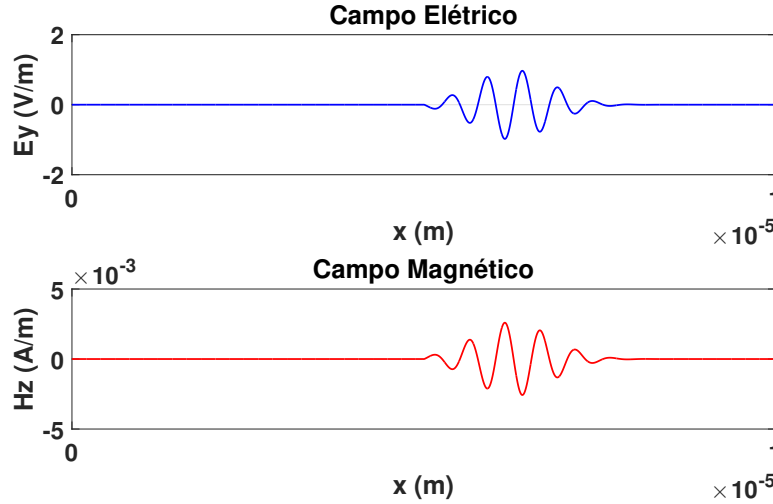


Figura 3.1: Propagação eletromagnética de um modo TEM_x no ar. O numero de elementos usados foi $N_e = 401$ e o grau do polinômio de Lagrange $p = 6$.

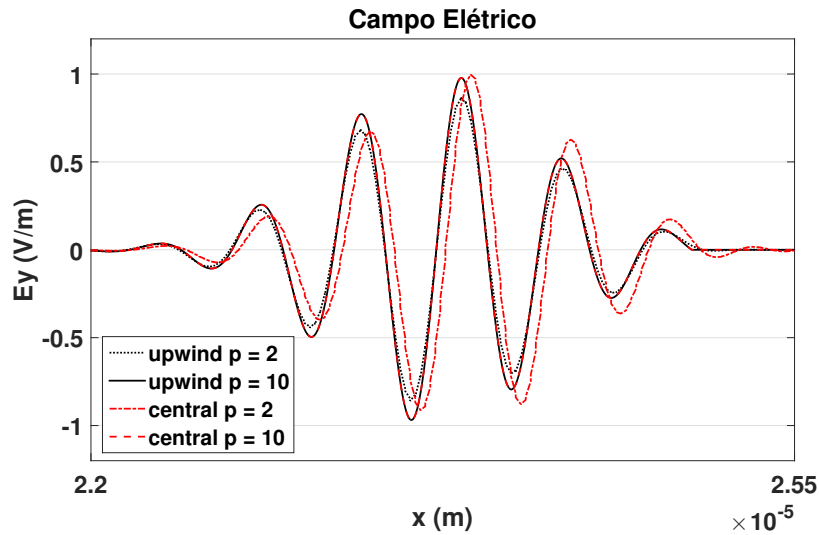


Figura 3.2: Comparação entre os fluxos central e *upwind*. O numero de elementos usados foi $N_e = 2001$.

3.5 Integração do Tempo

Diversos métodos podem ser usados na discretização do tempo. Estes métodos podem ser divididos em duas categorias: explícitos e implícitos. Métodos explícitos calculam o estado do sistema em um momento posterior usando a informação do estado no momento atual. Nos métodos implícitos, o estado em um momento posterior é encontrado com a solução de uma equação que envolve os estados no momento atual e posterior. Em outras palavras, nos métodos explícitos, $h(t + \Delta t) = f(h(t))$. Já nos métodos implícitos deve-se resolver uma equação de forma que: $Eq(h(t), h(t + \Delta t)) = 0$ Hirsch (2007). Técnicas de integração no tempo explícitas empregam métodos de diferenças finitas e são particularmente bem adaptadas para problemas dinâmicos de curta duração ou problemas de propagação de ondas. De acordo com Dokainish e Subbaraj (1989), quando a matriz de massa é diagonal não é necessário armazenar nenhuma matriz.

No DG, quando usadas funções de base ortogonais, a matriz de massa será diagonal. Além dos menores requisitos de armazenamento, o custo computacional por etapa de tempo é geralmente muito menor para métodos explícitos do que para métodos implícitos. Contudo, estes métodos são condicionalmente estáveis. Enquanto que nos métodos implícitos o passo de tempo afeta apenas a precisão da aproximação, nos métodos explícitos ele afeta a estabilidade do esquema. Normalmente, os métodos explícitos têm implementação mais simples que métodos implícitos. As características citadas, tornam os métodos explícitos adequados para serem usados em conjunto com o DG. Dentre eles, os mais usados são o *leapfrog* (LF) e o Runge-Kutta (RK).

Os métodos de integração do tempo de alta ordem podem ser divididos em duas categorias: métodos com múltiplos estágios e métodos com múltiplos passos. Para atingir alta ordem de precisão, métodos multi estágio usam estágios intermediários para o cálculo da solução. Considerando que: $\partial_t u = f(u, t)$ e que o objetivo seja evoluir $u(t_n) \rightarrow u(t_{n+1})$, sendo $t_{n+1} = t_n + \Delta t$. Em um método multi estágio com dois estágios, além de calcular $f(u_n)$, também é necessário calcular $f(u_{n+1/\ell})$. Sendo $t_{n+1/\ell}$ um valor de tempo intermediário entre t_n e t_{n+1} . Isso faz com que o custo computacional aumente em relação a um método com somente um estágio, porém eleva o grau de precisão para dois.

Nos métodos multi passo, para elevar a ordem de precisão, não são feitos cálculos adicionais em intervalos intermediários. Em vez disso, são consideradas soluções passadas no cálculo da solução futura. Ou seja, em um método de dois passos de ordem dois, no cálculo de u_{n+1} são usados os valores de u_{n-1} , $f(u_{n-1})$, u_n e $f(u_n)$. Os valores de u_{n-1} e $f(u_{n-1})$ podem ser armazenados, evitando terem que ser calculados a cada passo de tempo. Porém isso aumenta os requisitos de armazenamento.

3.5.1 Método *leapfrog* - LF

O *leapfrog* é muito usado no FDTD para solução de problemas eletromagnéticos. Uma de suas vantagens é o fato de ser simplético. Isto significa que a energia do sistema é conservada quando o tempo evolui. Este método é relativamente simples de ser implementado, e já é utilizado em esquemas DGTD. Normalmente é usado o LF2 (*second-order leapfrog*). Por exemplo, em [Alvarez et al. \(2015\)](#) é apresentada uma formulação DGTD que faz uso do LF2 aplicado à modelagem de antenas de banda larga. No *leapfrog* usado em [Alvarez et al. \(2015\)](#), as incógnitas são calculadas de maneira escalonada. O campo elétrico é calculado no instante $t_n = n\Delta t$, enquanto que o campo magnético é calculado no instante $t_{n+1/2} = (n + 1/2)\Delta t$. As derivadas temporais são aproximadas por diferenças finitas:

$$\partial_t H_n = \frac{H_{n+1/2} - H_{n-1/2}}{\Delta t} \quad \partial_t E_{n+1/2} = \frac{E_{n+1} - E_n}{\Delta t} \quad (3.56)$$

Com o objetivo de chegar a um esquema explícito, são feitas as seguintes aproximações: $H_n \approx H_{n-1/2}$ e $E_{n+1/2} \approx E_n$.

Apesar da maioria dos trabalhos encontrados na literatura usarem o LF2, existem versões deste esquema de ordens superiores. [Fahs \(2009\)](#) apresenta um esquema LFN de ordem arbitrária sendo utilizado em conjunto com o DG. O uso do LFN permite passos de tempo maiores que o LF2 e garante a alta convergência do esquema totalmente discreto quando são usadas funções de base de alta ordem. Porém, para um esquema de ordem N , o LFN necessita de $N/2$ mais memória e $N - 1$ vezes o número de operações algébricas que o LF2. Isto faz com que o LFN não seja muito utilizado ([Díaz Angulo \(2015\)](#)).

3.5.2 Método Runge-Kutta

Existem alguns tipos de esquemas Runge-Kutta. A principal vantagem do RK em relação ao LF é a facilidade de ajustar a ordem de integração no tempo com a ordem da função de base, p ([Álvarez González \(2014\)](#)). Isso garante a alta convergência do esquema totalmente discreto. Tipicamente, nestes esquemas é necessário calcular o campo em um número s de estágios auxiliares para que as derivadas temporais sejam efetuadas. Um dos esquemas mais usados é o LSRK (*low storage Runge Kutta*) apresentado em [Williamson \(1980\)](#). Este método é popular porque possui baixo requisito de armazenamento. Como será visto, o algoritmo LSRK é simples. Para isso considere que:

$$\partial_t \mathbf{q}_t = \mathcal{F}(\mathbf{q}_t, t) \quad (3.57)$$

Primeiramente, o tempo total deve ser dividido em intervalos $\Delta t = t_{n+1} - t_n$. Estes intervalos não precisam ser de tamanho uniforme. Então são seguidos os passos:

```

 $\mathbf{q}_0 = \mathbf{q}_t(t_n)$ 
Para  $i = 1$  até  $s$ 
{
 $\mathbf{p}_i = \mathcal{A}_i \mathbf{p}_{i-1} + \mathcal{F}(\mathbf{q}_{i-1}, t_n + \mathcal{C}_i \Delta t) \Delta t$ 
 $\mathbf{q}_i = \mathbf{q}_{i-1} + \mathcal{B}_i \mathbf{p}_i$ 
}
 $\mathbf{q}_t(t_{n+1}) = \mathbf{q}_s$ 

```

Algoritmo 1: LSRK

As constantes $\mathcal{A}_i$, $\mathcal{B}_i$ e $\mathcal{C}_i$ estão associadas à precisão e estabilidade do esquema. Uma tabela com os valores destas constantes para o LSRK de quarta ordem (LSRK4), é apresentada em [Álvarez González \(2014\)](#). Esta é a versão mais usada em conjunto com o DG de alta ordem. Além de permitir passos de tempo muito maiores que o LF2, o LSRK4 apresenta baixos erros de dispersão e dissipação. Porém, tem maior custo computacional que o LF2 e é dissipativo ([Díaz Angulo \(2015\)](#)).

3.5.3 Estabilidade Condicional

Os métodos LSRK e LF são esquemas explícitos e por isso, são condicionalmente estáveis. Portanto, estão sujeitos às condições CFL (Courant-Friedrichs-Lewy). A medida que o intervalo Δt excede seu limite máximo, a solução fica sujeita a um crescimento anormal que aumenta exponencialmente ([Busch et al. \(2011\)](#)). Existem diferentes valores do número CFL para diferentes tipos de esquemas de integração no tempo. Na literatura encontra-se trabalhos que visam desenvolver esquemas que preservem a precisão, mantendo um número CFL menos restritivo, por exemplo: [Sármány et al. \(2007\)](#). Dessa forma, é importante manter a estabilidade do esquema mantendo o intervalo Δt dentro do limite imposto pela condição CFL. Contudo, em se tratando de problemas envolvendo múltiplas escalas, com uma malha não estruturada, o passo de tempo deve se ajustar às dimensões menores e isso pode acarretar em um alto custo computacional. Uma alternativa é o uso do passo de tempo local.

3.5.4 Passo de Tempo Local - LTS

Como mostrado anteriormente no texto, o método DG é baseado em uma modelagem local. Cada elemento é tratado isoladamente. Isso favorece o uso de um passo de tempo local (*local time Step* - LTS) para cada elemento. Assim, o passo de tempo pode se ajustar às dimensões do elemento local. Um aspecto que promete ser um complicador, em termos de im-

plementação, é o fluxo numérico. Na implementação de um passo de tempo local, é necessário encontrar uma forma de tratar o fluxo numérico.

Já existem algumas técnicas que visam introduzir o passo de tempo local, tanto no LF2 quanto no LSRK4. Em [Angulo et al. \(2014\)](#) é apresentada uma eficiente estratégia LTS chamada *Causal Path Local Time Stepping* que pode ser aplicada a ambos os métodos.

3.5.4.1 Causal-Path Local Time Stepping - CPLTS

Em [Angulo et al. \(2014\)](#) é proposto um esquema que pode ser aplicado em dois métodos explícitos: o LF2 e o LSERK. Para possibilitar o uso desta estratégia o método de integração no tempo deve obedecer apenas dois requisitos:

- I - Ele tem que fornecer um estado totalmente definido $\mathbf{q}^k(t)$ para cada elemento.
- II - O próximo estado $\mathbf{q}^k(t + \Delta t)$ pode ser explicitamente calculado a partir da vizinhança do elemento k .

Neste método, o primeiro passo é dividir os elementos em diferentes grupos denominados de camadas. A classificação se dá de acordo com o seu passo de tempo Δt^m . Um elemento k pertencerá a uma camada $m = [0, \dots, N_m - 1]$, se o máximo passo de tempo do elemento Δt_k for:

$$\Delta t^m \leq \Delta t_k < \Delta t^{m+1} \quad (3.58)$$

Isso significa que os elementos com Δt^m 's menores pertencerão a camada com $m = 0$, enquanto que os com Δt^m 's maiores pertencerão a camada com $m = N_m - 1$. As camadas são ordenadas de forma que $\Delta t^m < \Delta t^{m+1}$. Em cada camada m existe um número diferente de etapas na marcha onde são computados os campos em cada elemento. As camadas mais baixas possuem um número maior de etapas. Vizinhos que não residem na mesma camada devem residir em camadas adjacentes. Neste método são calculados estágios adicionais em camadas mais altas e passadas para as mais baixas para que o fluxo numérico seja calculado. Assim, não são necessárias interpolações. Em [Angulo et al. \(2014\)](#), o CPLTS é aplicado ao LF2 e ao LSRK4.

3.5.4.2 LF2-CPLTS

Na marcha do LF2 em um passo de tempo, possui somente uma etapa, mas é possível criar N_s etapas intermediárias. Nas camadas mais altas do LF2-CPTLS são calculados etapas intermediárias nos elementos que fazem fronteira com elementos da camada mais baixa, a fim

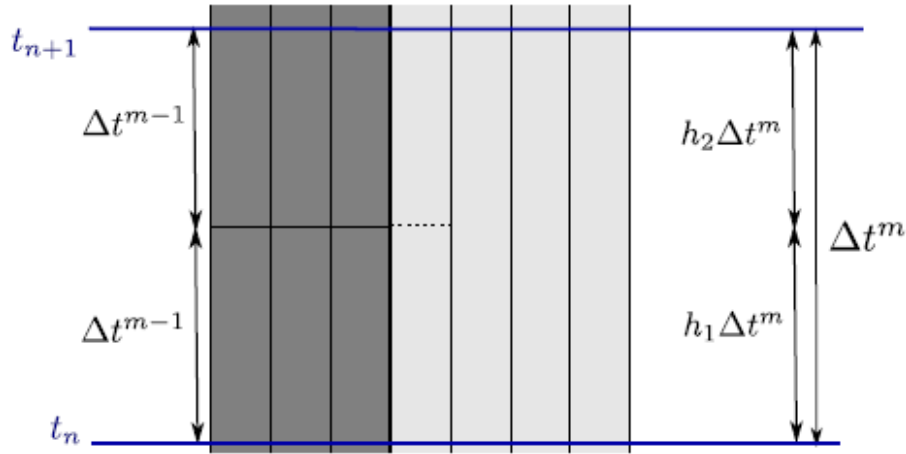


Figura 3.3: Passo local no LF2-CPTLS com $h_1 = 1/2$ e $h_2 = 1/2$. [Angulo et al. \(2014\)](#)

de prover a informação necessária para calcular o fluxo numérico. O passo de tempo em cada etapa intermediária é dado por: $h_i \Delta t^{m,i} = \Delta t^{m-1}$, sendo $h_i > 0$ e com a restrição $\sum_{i=1}^{N_s} h_i = 1$. A Figura 3.3 encontrada em [Angulo et al. \(2014\)](#) ajuda a entender melhor o esquema. Nela é ilustrado de forma esquemática como funciona o LF2-CPTLS no caso particular onde $h_1 = h_2 = 1/2$. As linhas verticais representam as fronteiras entre os elementos. Em t_n todos os valores de campo são conhecidos. Na região mais escura, que representa a camada $m - 1$ e possui passo de tempo Δt^{m-1} , é necessário usar o LF2 duas vezes para chegar aos valores de campo em t_{n+1} . O algoritmo LF2-CPTLS calcula o LF2 duas vezes também nos elementos vizinhos aos da camada $m - 1$. Estes elementos são representados em 3.3 por uma linha horizontal tracejada.

3.5.4.3 LSRK4-CPTLS

Quando o método CPLTS é aplicado ao esquema LSRK4, os intervalos entre as etapas, $h_i \Delta t^{m,i}$, não precisam ser distribuídos uniformemente no tempo. Dessa forma, os etapas intermediárias das camadas mais baixas podem ocorrer em intervalos de tempo variáveis. Os valores de Δt^{m-1} e Δt^m podem ser escolhidos de forma a satisfazer os critérios de estabilidade dos elementos de cada camada. Isto traz uma maior flexibilidade ao esquema e os passos de tempo de cada camada prometem ser mais ajustados às dimensões de seus elementos. Contudo, essa aproximação só traz um ganho computacional real se a camada superior tiver pelo menos $5/3$ vezes o número de elementos na camada inferior. Caso contrário é melhor que os etapas intermediárias sejam uniformemente distribuídos, de forma que: $\Delta t^{m-1} = \Delta t^m / N_s$.

O método proposto por [Angulo et al. \(2014\)](#) se mostra eficiente e pode ser aplicado em dois métodos de integração do tempo, o LF2 (LF2-CPLTS) e métodos da família Runge Kutta (LSRK-CPLTS). Apesar dos autores demonstrarem sua eficiência, esses métodos apresentam sérias limitações. O LF2 possui baixa ordem de precisão, e por isso, o passo de tempo se torna muito restritivo quando um erro baixo é desejado.

O LESRK4-CPLTS é de quarta ordem e possui cinco estágios. A fim de satisfazer o fluxo numérico, é necessário que as camadas vizinhas a outras com passo de tempo menor, possuam pelo menos quatro subcamadas de elementos. De forma geral, quando usado um método com s estágios como base da estratégia CPLTS, as subcamadas devem conter no mínimo $s - 1$ elementos para que o fluxo possa ser satisfeito. Consequentemente, esse esquema pode ter dificuldade em se adaptar a determinadas discretizações espaciais. Além disso, os métodos com múltiplos estágios requerem um grande número de operações matemáticas a cada passo de tempo. Com o aumento da ordem de precisão, o número de estágios cresce fazendo com que o número de operações matemáticas necessárias a cada passo de tempo aumente de forma significativa.

A estratégia LTS que será apresentada aqui (método multiclasse) pode ser aplicada em qualquer método multi-passo. Porém, aqui será aplicada especificamente à família *linear strong stability preserving multistep methods* (SSPMS). Dessa forma, o método proposto pode reunir as vantagens do LF2-CPLTS e do LESRK-CPLTS. Como no LF2-CPLAT, no método multiclasse só é necessário uma subcamada de elementos em uma classe vizinha para garantir o fluxo numérico. Isso facilita a implementação e torna a estratégia mais facilmente adaptável às diversas discretizações espaciais. Além disso, com o uso do multi-passo é possível criar um método de alta ordem de precisão, sem aumentar significativamente o número de operações matemáticas a serem realizadas a cada passo de tempo.

3.6 Método multi passo linear com forte estabilidade - SSPMS

Diferentemente do LESRK, no método SSPMS não existem estágios intermediários no cálculo da solução. Isso reduz o número de operações e traz a possibilidade de ganho computacional. A alta ordem é alcançada usando soluções passadas, levando a um aumento nos requisitos de armazenamento. Assim, neste trabalho, propomos um *trade-off* entre armazenamento e custo computacional.

Os métodos de integração de tempo podem ser considerados como ODE *solvers*. A priori, considerando métodos de primeira ordem para resolver o problema:

$$\mathbf{q}'(t) = \mathcal{H}(\mathbf{q}, t); \quad \mathbf{q}(t_0) = \mathbf{q}_0. \quad (3.59)$$

Um método simples de primeira ordem para resolver (3.59) é o Euler *forward*. O método Euler *forward* é explícito e pode ser derivado, entre outras formas, usando uma expansão em série de Taylor. Considerando um intervalo de tempo Δt :

$$\mathbf{q}(t_0 + \Delta t) = \mathbf{q}(t_0) + \Delta t \mathbf{q}'(t_0) + O(\Delta t^2). \quad (3.60)$$

Em (3.59) pode ser visto que $\mathbf{q}'(t) = \mathcal{H}(\mathbf{q}, t)$. Portanto, aplicando o teorema fundamental do cálculo:

$$\mathbf{q}(t_0 + \Delta t) - \mathbf{q}(t_0) = \int_{t_0}^{t_0 + \Delta t} \mathcal{H}(\mathbf{q}, t) dt. \quad (3.61)$$

Sendo Δt pequeno o suficiente, a integral no lado direito de (3.61) pode ser aproximada por $\Delta t \mathcal{H}(\mathbf{q}_0, t_0)$. Dessa forma, o método Euler *forward*, para o intervalo $t_0 \rightarrow t_0 + \Delta t$, é dado por:

$$\mathbf{q}(t_0 + \Delta t) = \mathbf{q}_0 + \Delta t \mathcal{H}(\mathbf{q}_0, t_0). \quad (3.62)$$

Isso pode ser generalizado para qualquer intervalo. Alterando a notação e reescrevendo (3.62) para um intervalo genérico $t_n \rightarrow t_{n+1}$, onde $t_{n+1} = t_n + \Delta t$:

$$\mathbf{q}_{n+1} = \mathbf{q}_n + \Delta t \mathcal{H}(\mathbf{q}_n, t_n). \quad (3.63)$$

Considerando os termos de ordem superiores da série de Taylor, é possível chegar aos esquemas lineares de múltiplos passos, chamados Adams-Bashforth (Booth (1954)). Nos métodos Runge-Kutta com ordem de precisão maior que quatro (ordem $o > 4$), o número de estágios pode crescer significativamente ($s > o$), aumentando o custo computacional. Ao contrário do RK, em métodos de múltiplos passos, o crescimento da ordem de precisão não aumenta significativamente o número de operações matemáticas em cada passo. Portanto, a ordem da integração temporal pode ser ajustada pela ordem da discretização espacial gerando um aumento menor no custo computacional.

O método Euler *forward* é um esquema TVD (*total variation diminishing*). Portanto, ele tem a propriedade SSP (*strong stability preserving*) (Bresten *et al.* (2017)). Isto significa que para uma dada norma $\|\cdot\|$, tem-se: $\|\mathbf{q}_n + \Delta t \mathcal{H}(\mathbf{q}_n, t_n)\| \leq \|\mathbf{q}_n\|$ para $0 \leq \Delta t \leq \Delta t_{FE}$. Sendo Δt_{FE} o passo tempo no método Euler *forward*. A propriedade SSP garante a monotonicidade da integração do tempo. O passo de tempo de referência para o DG pode ser calculado por Hesthaven e Warburton (2002a):

$$\Delta t_{FE} = \frac{2}{3} r_G \min\left(\frac{\mathbf{r}_D}{|c|}\right). \quad (3.64)$$

Em (3.64), $|c|$ é o módulo da velocidade da onda, $\mathbf{r}_D$ é um vetor que contém a razão entre a área e o semiperímetro de cada elemento e r_G é a menor distância entre dois pontos no elemento de

referência unidimensional, $[-1, 1]$, com a mesma ordem da aproximação espacial. Os métodos lineares multipasso com SSP podem ser construídos tomando uma combinação linear do Euler *forward* (Ruuth e Hundsdorfer (2005)).

$$\mathbf{q}_{n+1} = \sum_{i=1}^s (\alpha_i \mathbf{q}_{n+1-i} + \beta_i \Delta t \mathcal{H}(\mathbf{q}_{n+1-i}, t_{n+1-i})). \quad (3.65)$$

Para manter a consistência, é necessário que $\sum_i \alpha_i = 1$. O intervalo de tempo é uma função do coeficiente SSP: $\Delta t = \mathcal{C}(\alpha_i, \beta_i) \Delta t_{FE}$ com (Gottlieb *et al.* (2011)):

$$\mathcal{C} = \min \left(\frac{\beta_i}{\alpha_i} \right), \quad \text{if } \alpha_i > 0 \text{ for all } i \quad (3.66)$$

Portanto, em um método SSP de alta ordem: $\|\mathbf{q}_{n+1}\| \leq \|\mathbf{q}_n\|$ para $0 \leq \Delta t \leq \mathcal{C} \Delta t_{FE}$. Uma descrição detalhada das características dos métodos SSP pode ser encontrada em Gottlieb *et al.* (2011).

3.6.1 Método multi classe *mC*-SSPMS

Nossa estratégia é simples, não aumenta os requisitos de armazenamento e reduz o número de operações matemáticas a serem realizadas em cada passo de tempo. No esquema proposto, os elementos serão classificados por seus respectivos passos de tempo locais. Primeiro, o passo de tempo local Δt_k de cada elemento local é calculado. Em seguida, são agrupados os elementos em m classes de acordo com o Δt_k . O Δt_k menos restritivo é tomado como referência, ou seja, $\Delta t_{ref} = \max(\Delta t_k)$. Então, os passos de tempo em cada classe j são dados por: $\Delta t_j = \Delta t_{ref} / 2^{\kappa_j}$, $\kappa_j \in \mathbb{N}$. Dessa forma, a sincronização é feita facilmente. Na primeira classe, $\kappa_1 = 0$, nas outras classes κ_j dependerá da diferença de tamanho dos elementos da malha.

Antes de continuar, há uma questão importante que deve ser considerada: o fluxo numérico deve ser satisfeito nas interfaces de cada elemento. Elementos de uma classe que têm vizinhos de outra classe, com um intervalo de tempo maior, precisarão de valores nodais em instantes que não são calculados no tempo de marcha. A fim de satisfazer o fluxo numérico nessas interfaces, são realizadas interpolações da solução, nos nós da aresta vizinha do elemento de classe com o maior passo de tempo, nos instantes requeridos. Cálculos adicionais não serão necessários para satisfazer o fluxo numérico. Vários tipos de interpolação podem ser utilizadas. A escolha deve ser feita com base na precisão e no custo computacional. Por exemplo, uma interpolação linear pode ser adequada quando a discrepância entre os passos de tempo de classes adjacentes não é muito grande. No entanto, quando o salto de tempo de uma classe para outra é muito grande, é recomendado o uso interpolações polinomiais com ordens mais altas

para melhorar a precisão. É importante ressaltar que, com o uso da interpolação, não é possível garantir que o esquema multi classe conserve a propriedade SSP.

A marcha no tempo, em cada classe, pode ser realizada em paralelo. No entanto, nesse trabalho, foi desenvolvida uma implementação sequencial eficiente, que pode produzir uma aceleração significativa em grandes sistemas, com grandes janelas tempo e malhas contendo múltiplas escalas. Dessa forma, a solução pode ser representada como:

$$\mathbf{q}_{n+1} = \sum_{j=1}^m A_j \mathbf{q}_{n+1} = \sum_{j=1}^m \mathbf{q}_{n+1}^j. \quad (3.67)$$

Em (3.67), A_j são matrizes esparsas com elementos iguais à unidade em algumas entradas das diagonais principais. Essas entradas referem-se aos nós contidos na classe j . A solução dos nós da classe j é representada como $\mathbf{q}_{n+1}^j$. Na classe $j = 1$, quando o tempo passa de t_n para t_{n+1} , o cálculo é feito em uma etapa. Nas classes $j > 1$, é necessário executar cálculos intermediários, ou sub-passos, até atingir t_{n+1} . O número de sub-passos dependerá de κ_j .

É possível entender melhor o processo observando o seguinte exemplo. Considere três classes com $\kappa_1 = 0$, $\kappa_2 = 1$ e $\kappa_3 = 2$. Portanto, $\Delta t_2 = \Delta t_1/2$ e $\Delta t_3 = \Delta t_2/2 = \Delta t_1/4$. Além disso, considere que em t_n todos os valores nodais são conhecidos. Em $t_{n+1/4}$, $\mathbf{q}_{n+1/4}^3$ é atualizado. No momento $t_{n+1/2}$, $\mathbf{q}_{n+1/2}^3$ e $\mathbf{q}_{n+1/2}^2$ são atualizados. Em $t_{n+3/4}$, $\mathbf{q}_{n+3/4}^3$ é atualizado. Finalmente, no instante t_{n+1} , $\mathbf{q}_{n+1}^3$, $\mathbf{q}_{n+1}^2$ e $\mathbf{q}_{n+1}^1$ são atualizados. Este processo é ilustrado em Fig. 3.4.

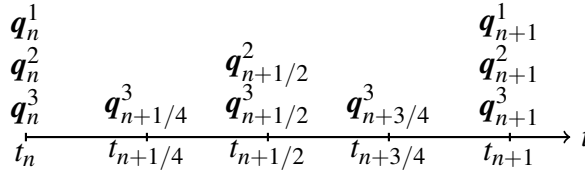


Figura 3.4: Evolução de classes.

O número de sub-passos de uma classe j é: $n_s^j = 2^{\kappa_j} - 1$. Não deve ser esquecido que os métodos multi-passo utilizam soluções anteriores para calcular a solução na próxima etapa. No entanto, isso é necessário apenas no primeiro instante da marcha do tempo. Depois disso, as soluções podem ser armazenadas e usadas nos cálculos de soluções futuras. Em geral, um sub-passo em uma classe j com número n_s de sub-passos pode ser calculado como:

$$\mathbf{q}_{n+\ell/2^{\kappa_j}}^j = \sum_{i=1}^s (\alpha_i \mathbf{q}_{n+(\ell-i)/2^{\kappa_j}}^j + \beta_i \Delta t_j \mathcal{H}_j(\mathbf{q}_{n+(\ell-i)/2^{\kappa_j}}^j, t_{n+(\ell-i)/2^{\kappa_j}})) \quad (3.68)$$

com ℓ variando entre 1 e $n_s + 1$. É importante destacar que $\mathcal{H}_j$ opera somente em elementos contidos na classe j . No entanto, para calcular o fluxo numérico em $\mathcal{H}_j$, é necessário fornecer

informações sobre os elementos vizinhos. Incluindo valores interpolados. O procedimento para um passo da marcha no tempo é demonstrado no seguinte pseudocódigo:

```

for  $j = 1$  until  $m$  do
  for  $\ell = 1$  until  $n_s + 1$  do
     $S \leftarrow 0$ 
    for  $i = 1$  until  $s$  do
      if  $i = 1$  then
        Compute the contribution to:
         $\mathcal{H}_j(\mathbf{q}_{n+(\ell-i)/2^{\kappa_j}}^j, t_{n+(\ell-i)/2^{\kappa_j}})$ 
      else
        load  $\mathcal{H}_j(\mathbf{q}_{n+(\ell-i)/2^{\kappa_j}}^j, t_{n+(\ell-i)/2^{\kappa_j}})$ 
      end if
       $S \leftarrow S + \alpha_i \mathbf{q}_{n+(\ell-i)/2^{\kappa_j}}^j + \beta_i \Delta t_j \mathcal{H}_j$ 
    end for
     $\mathbf{q}_{n+\ell/2^{\kappa_j}}^j \leftarrow S$ 
  end for
end for

```

CAPÍTULO 4

Testes Numéricos

Nesse capítulo serão apresentados alguns teste numéricos realizados com o objetivo de testar e validar a estratégia multi classe que foi descrita anteriormente.

4.1 Equação de Advecção

A análise de complexidade indica que o método multi classe pode trazer uma melhoria significativa na eficiência computacional das simulações, onde os sistemas são grandes e precisam de uma grande janela de tempo. Entretanto, é importante testar o método experimentalmente para observar se um ganho real é alcançado. Para controlar facilmente o número de elementos e a relação entre eles, primeiro foram realizados experimentos com a discretização espacial unidimensional da equação de Advecção. Nos cenários criados para os experimentos, o número de elementos em cada classe variou, assim como a relação entre o tamanho dos elementos. Considerando que o número de elementos em uma classe é dado por N_{el}^j , e que o tamanho dos elementos por h_j , os experimentos foram realizados da seguinte forma:

- $h_1 = 2h_2$
- $h_1 = 4h_2$
- $h_1 = 8h_2$

Com o número dos elementos variando da forma: $N_{el}^1 = N_{el}^2$ e $N_{el}^1 = 3N_{el}^2$, onde $N_{el} = N_{el}^1 + N_{el}^2$.

Nestas experiências, foram comparados três métodos. O método padrão Runge Kutta de quarta ordem (Gottlieb e Shu (1998)), o SSPMS de ordem três e seis passos, apresentado em Ketcheson (2009). Além disso, foi construído um esquema de duas classes com base no método multi passo (2C-SSPMS(6,3)). O SSPMS(6,3) tem um coeficiente SSP $\mathcal{C} = 0.52$. Os coeficientes α_i e β_i são:

$$\begin{aligned}
\alpha_1 &= 0.850708871672579 & \beta_1 &= 1.459638436015276 \\
\alpha_5 &= 0.030664864534383 & \beta_5 &= 0.052614491749200 \\
\alpha_6 &= 0.118626263793039 & \beta_6 &= 0.203537849338252
\end{aligned}$$

O tempo de execução, de todos os métodos em todos os cenários, foi medido com o objetivo de compará-los. Os resultados são mostrados na Figura 4.1. É interessante notar a similaridade entre os resultados do RK4 e o SSPMS(6,3) em termos do tempo de execução. As curvas de tempo de execução desses métodos são quase iguais em todos os cenários.

Observando a Figura 4.1 percebe-se que, em geral, a estratégia multiclasse trouxe ganho computacional ao esquema. A Figura 4.1 (a) mostra o tempo para o sistema com $h_1 = 2h_2$ e $N_{el}^1 = N_{el}^2$. Percebe-se que não houve ganho significativo de eficiência neste cenário. Figura 4.1 (b) mostra o tempo usado para integrar o sistema com $h_1 = 2h_2$, mas com $N_{el}^1 = 3N_{el}^2$. Ou seja, o número de elementos na classe 1, com $dt_1 = 2dt_2$, é três vezes maior que o número de elementos na classe 2. Aqui nota-se uma pequena melhoria de eficiência. O terceiro cenário é apresentado na Figura 4.1 (c), quando $h_1 = 4h_2 \Rightarrow dt_1 = 4dt_2$. Os elementos foram divididos igualmente entre as classes. Novamente nota-se uma ligeira melhoria na eficiência computacional, mas menor do que no cenário anterior mostrado na Figura 4.1 (b). No entanto, quando $h_1 = 4h_2$ e a relação entre o número de elementos se tornou $N_{el}^1 = 3N_{el}^2$, há uma grande melhoria na eficiência, como é possível observar na Figura 4.1 (d). O cenário mostrado na Figura 4.1 (e) refere-se à situação em que $h_1 = 8h_2$ e, consequentemente, $dt_1 = 8dt_2$, com elementos igualmente divididos entre classes. Embora a inclinação da curva de tempo do esquema LTS sugira que: quando o número de elementos tende a crescer, o método multiclasse trará um ganho computacional. No intervalo em que foi simulado isso não aconteceu. O último cenário, com $h_1 = 8h_2$ e $N_{el}^1 = 3N_{el}^2$, tem seu tempo de execução mostrado na Figura 4.1 (f). Nesse cenário, o esquema LTS produz uma melhoria significativa na eficiência computacional da integração de tempo.

Para manter a estabilidade numérica, na implementação da estratégia LTS, o passo de tempo global deve ser reduzido por um fator. Em outras palavras, $\Delta t_{mC} = C_r \Delta t_0$, onde $0 \leq C_r \leq 1$. Na Tabela 4.1 são apresentados o fator de redução de todos os cenários. Quando a relação entre o número de elementos é $N_{el}^1 = N_{el}^2$ a diminuição do intervalo de tempo global deve ser feita progressivamente enquanto é aumentada a discrepância entre h_1 e h_2 . No entanto, quando $N_{el}^1 = 3N_{el}^2$ o fator em que o passo de tempo global deve ser reduzido, é mantido constante. Isso indica que a distribuição dos elementos nas classes pode influenciar a estabilidade e a eficiência do esquema. Dessa forma, a estratégia LTS trará uma melhoria mais significativa para problemas em que o número de elementos na classe superior seja maior que o número de elementos de classes mais baixas.

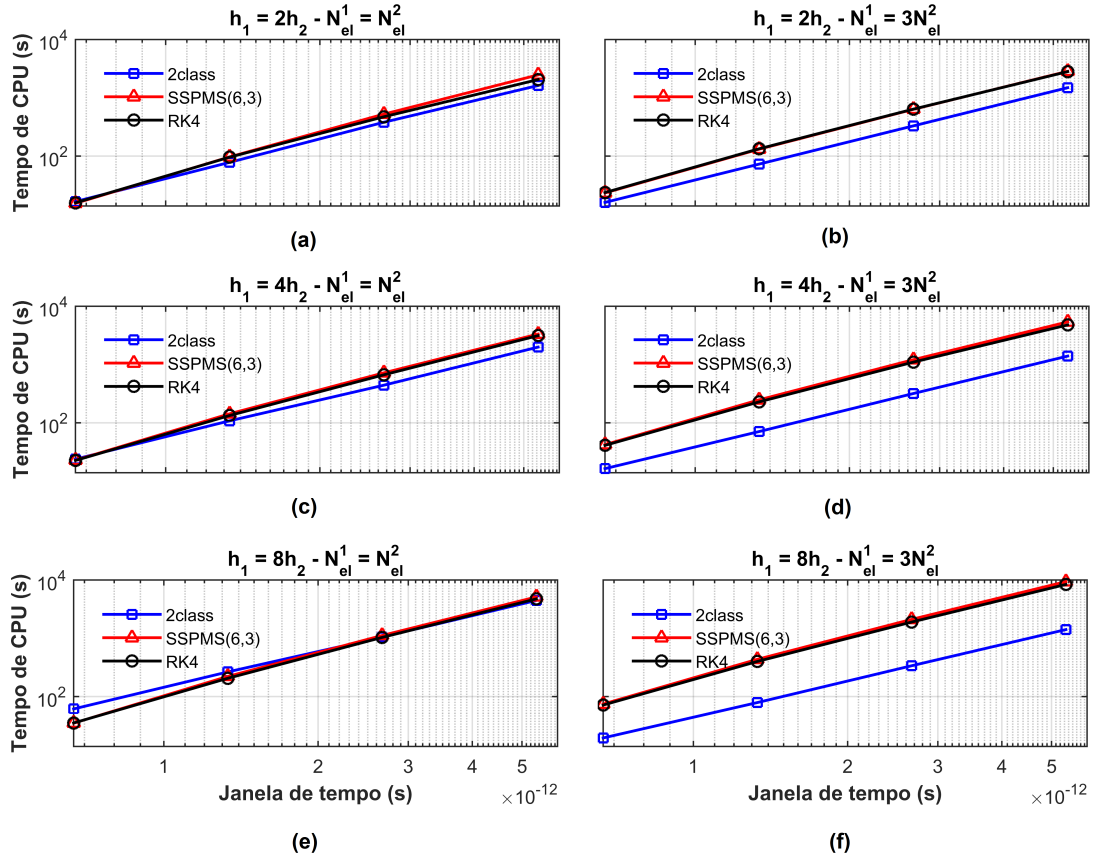


Figura 4.1: Equação de Advecção: testes experimentais.

Tabela 4.1: Equação de Advecção: fator de redução C_r

	$h_1 = 2h_2$	$h_1 = 4h_2$	$h_1 = 8h_2$
$N_1 = N_2$	0.80	0.53	0.26
$N_1 = 3N_2$	0.80	0.80	0.80

O erro relativo em cada cenário é apresentado na Figura 4.2. Em todos os cenários, observa-se que o erro relativo máximo é menor que 4×10^{-4} . Outro fato é que o SSPMS(6,3), em geral, demonstrou mais acurácia que os esquemas RK4 e 2C-SSPMS(6,3). Nota-se que em todos os cenários o erro tem uma pequena variação na janela de tempo. Esse comportamento ocorre porque quando a janela de tempo é aumentada, o domínio de espaço L também é expandido. No entanto, o tamanho dos elementos é mantido na mesma proporção. Em outras palavras, L/h_j é constante independentemente da janela de tempo. Nos cenários em que $N_{el}^1 = N_{el}^2$, o erro do 2C-SSPMS(6,3) e do SSPMS convencional, é quase o mesmo. O erro do esquema RK4 começa maior do que os outros, mas diminui com o aumento da janela de tempo e se torna quase igual aos outros. No caso em que $N_{el}^1 = 3N_{el}^2$ com $h_1 = 2h_2$ e $h_1 = 4h_2$, o erro do 2C-SSPMS(6,3) é quase o mesmo que SSPMS e RK convencionais. Onde $N_{el}^1 = 3N_{el}^2$

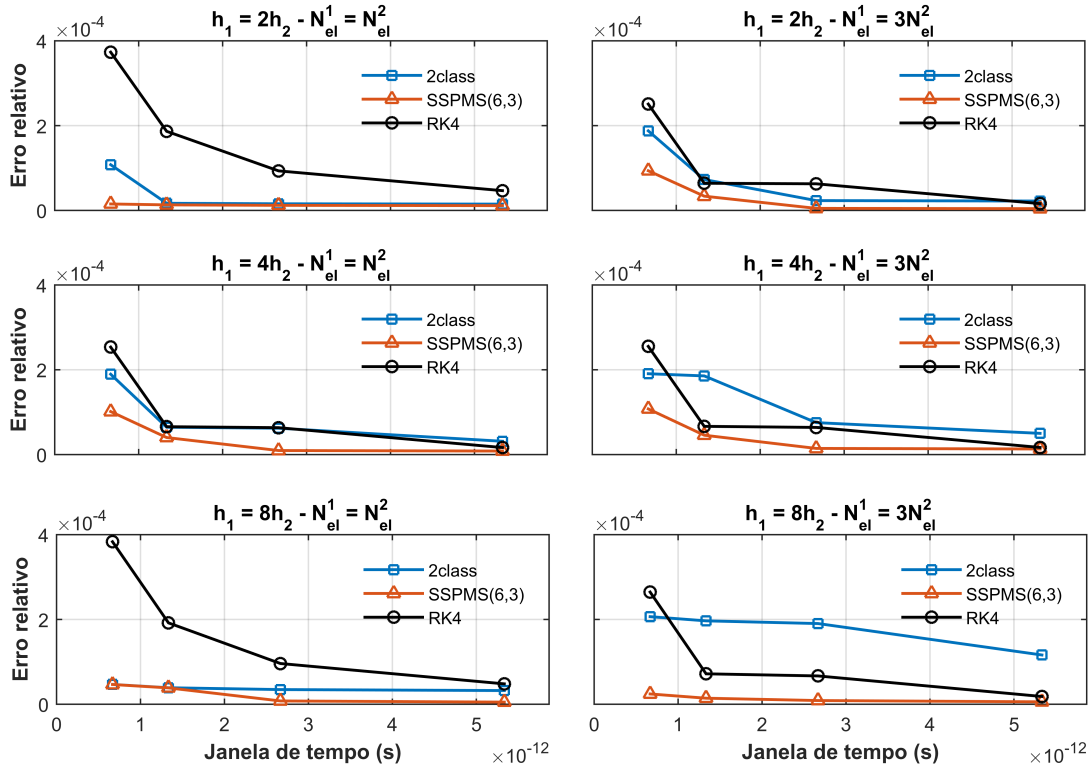


Figura 4.2: Equação de Advecção: erro relativo.

e $h_1 = 8h_2$, o erro 2C-SSPMS(6,3) se torna um pouco maior que as versões convencionais. Nesse caso, há uma troca que envolve precisão e eficiência computacional. Isso ocorre porque quando a janela de tempo se expande, a diferença de erro entre os esquemas 2C-SSPMS(6,3) e os convencionais é quase mantida, por outro lado, a diferença do tempo computacional do convencional aumenta exponencialmente. Além disso, embora exista uma diferença, o erro ainda é da mesma ordem. Portanto, em uma malha com o mesmo refinamento, é possível aumentar a eficiência computacional com o custo de menor precisão na solução.

4.2 Propagação Eletromagnética

Agora serão resolvidos problemas de propagação de ondas eletromagnéticas, a fim de verificar a eficiência e a precisão do *mC*-SSPMS(6,3).

4.2.1 Teste 1: Propagação de um pulso gaussiano

Nesse teste, um modo TM_z (H_x, H_y, E_z) é carregado usando a condição inicial do campo elétrico no domínio. A forma de onda do pulso gaussiano é:

$$E_z(x, y, t = 0) = e^{-\frac{x^2+y^2}{4\sigma^2}} \quad (4.1)$$

onde a largura espacial do pulso é $\sigma = 0.1$ m. Para truncar o domínio computacional, a condição de contorno Silver-Muller de primeira ordem $\mathbf{n} \times \mathbf{E} - Z\mathbf{n} \times (\mathbf{H} \times \mathbf{n}) = 0$ é aplicada no limite do círculo de raio $r = 5,0$ m. Foi utilizada uma malha triangular não uniforme, constituída por 2721 vértices e 5312 triângulos, como mostrado na Figura 4.3. A razão entre as arestas maior e menor é de 16.13 e, neste caso, o intervalo que contém todos os passos de tempo locais é dado por $(2/3c)r_G[1.26, 18.13]$ s. A propagação do pulso gaussiano é ilustrada na Figura 4.7, que apresenta quatro soluções em instantes diferentes.

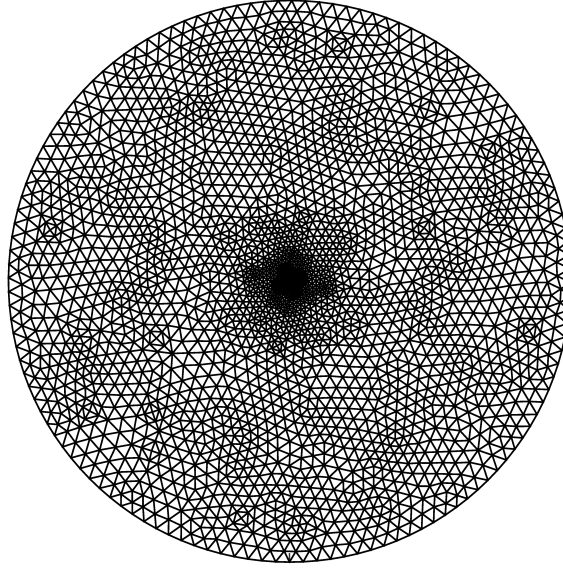


Figura 4.3: Malha multiescala triangular para resolver o problema de propagação de pulso.

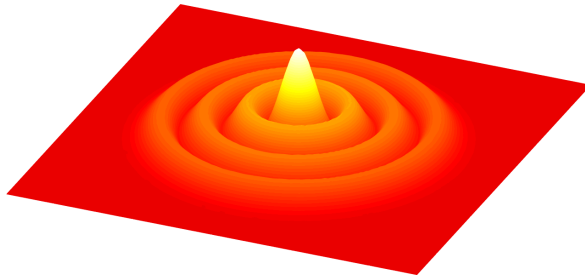


Figura 4.4: Teste 1: propagação do pulso gaussiano com quatro soluções em instantes diferentes.

O objetivo do teste 1 é comparar a eficiência computacional (tempo de CPU) e a precisão do mC -SSPMS(6,3) e o SSPMS(6,3) padrão e o LF2-CPLTS. Dessa forma, na simulação usando

Tabela 4.2: Teste 1: Divisão de elementos entre classes

Elementos por classe	2C-SSPMS(6,3)	3C-SSPMS(6,3)	4C-SSPMS(6,3)	5C-SSPMS(6,3)
Classe 1	4972	4668	4152	2992
Classe 2	340	392	516	1160
Classe 3		252	392	516
Classe 4			252	392
Classe 5				252

Tabela 4.3: Teste 1: eficiência e precisão de *mC*-SSPMS(6,3), SSPMS(6,3) e RK4

N	2C-SSPMS(6,3)		3C-SSPMS(6,3)		4C-SSPMS(6,3)		5C-SSPMS(6,3)		SSPMS(6,3)		RK4	
	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r
1	2.140	0.1692	1.895	0.1695	1.554	0.1718	1.593	0.1830	3.970	0.2470	6.803	0.2442
2	5.949	0.0579	4.882	0.0596	4.002	0.0597	4.770	0.0623	10.079	0.0791	16.994	0.0789
3	13.495	0.0442	11.062	0.0447	9.013	0.0434	11.294	0.0451	21.384	0.0438	36.706	0.0437
4	27.092	0.0264	21.879	0.0267	17.836	0.0253	23.625	0.0264	42.696	0.0260	73.069	0.0259
5	45.880	0.0166	38.416	0.0167	29.321	0.0156	38.851	0.0165	75.315	0.0166	129.889	0.0166

o *mC*-SSPMS(6,3), *m* variou de: 2, 3, 4, 5. O número de elementos em cada classe, para todos os valores de *m*, é mostrado na Tabela 4.2. Além disso a estratégia proposta também foi comparada com o Runge-Kutta de quarta ordem (RK4) padrão, que é muito utilizado na literatura.

A precisão de cada esquema foi medida, em cada janela de tempo, calculando o erro L2 como:

$$e_r = \int_{\Omega} \sqrt{\left(\frac{\mathbf{E}_h - \mathbf{E}}{\mathbf{E}}\right)^2} d\Omega \quad (4.2)$$

onde $\mathbf{E}_h$ é a solução DGTD e $\mathbf{E}$ é a solução de referência. Nas simulações, foi utilizado o fluxo numérico *upwind* e a interpolação linear. O tempo de execução e o erro de cada esquema são apresentados na Tabela 4.3. Aqui, o erro é calculado em $t = 1,3 ns$, quando a simulação é interrompida.

O primeiro fato a ser destacado é o ganho de eficiência do SSPMS(6,3) padrão em relação ao RK4. O tempo de execução da marcha com o método SSPMS(6,3) ficou em torno de 42% menor do que com o RK4. Isso ocorreu devido aos estágios adicionais no RK4. Mesmo com um intervalo de tempo de cerca de metade, mas tendo apenas um estágio e não quatro como o RK4, era esperado um ganho de cerca de 50%. De acordo com os resultados da Tabela 4.3, em geral, a divisão multi classe apresentou melhoria na eficiência computacional.

Nesse teste, todas as versões *mC*-SSPMS(6,3) apresentaram *speedup* em relação à versão padrão. O 4C-SSPMS(6,3) apresentou o melhor desempenho, sendo em média 2.48 vezes mais rápido que o SSPMS(6,3) padrão, e 4.24 vezes mais rápido que o RK4.

Tabela 4.4: Teste 1: Fator C_r para o mC -SSPMS(6,3) com interpolação

N	2C-SSPMS	3C-SSPMS	4C-SSPMS	5C-SSPMS
1	0.98	0.81	0.81	0.75
2	0.83	0.65	0.65	0.51
3	0.83	0.65	0.65	0.50
4	0.63	0.65	0.65	0.46
5	0.61	0.58	0.53	0.44

Como esperado, o desempenho do esquema multiclasse é afetado pelo número de classes, bem como pelo número de elementos que elas contém. Por exemplo, ao aumentar o número de classes, o processo de interpolação em cada etapa de tempo pode degradar a estabilidade, e introduzir uma sobrecarga de tempo computacional. Portanto, é necessário desenvolver uma estratégia para encontrar o número de classes para obter um bom desempenho em termos de esforço computacional. Uma maneira de resolver esse problema é agrupar os elementos de modo que a discrepância do passo de tempo médio entre as classes seja reduzida. Para capturar essas informações, foi definido um parâmetro de classe dado por:

$$r_j = \frac{1}{K_j} \sum \frac{\Delta t_{k_j}}{\min(h_{k_j})}. \quad (4.3)$$

K_j é o número de elementos na classe j , Δt_{k_j} é o passo de tempo local de k_j e $\min(h_{k_j})$ é o tamanho da menor aresta do elemento k_j .

Assim, propomos o seguinte passo de pré-processamento: para uma dada malha suportando até M classes. Primeiro, os parâmetros $r_j, j = 1, \dots, m$, são calculados para cada classe $m = 2, \dots, M$. Os pontos gerados (j, r_j) são marcados em um plano e a inclinação, ϕ_m , do melhor ajuste de primeira ordem é um parâmetro de não uniformidade que indica a variação de Δt entre classes. Figura 4.5 ilustra a aplicação deste processo. Observa-se uma saturação no parâmetro ϕ_m quando aumentado o número de classes. A saturação indica que a uniformização do Δt de classes tende a um limite. Portanto, o aumento arbitrário no número de classes não garante melhoria no tempo de CPU. Assim, um número adequado de classes estará próximo do regime de saturação. No Teste 1, $m = 4$ é o primeiro valor antes do regime de saturação.

Para garantir a estabilidade e precisão, o passo de tempo deve ser reduzido por um fator C_r (Tabela 4.4). Assim, o passo de tempo efetivo usado em cada simulação é dado por $\Delta t_j = C_r \Delta t$. Essa redução aumenta quando aumentado o número de classes. Provavelmente isso se deve ao aumento nas interpolações. Parece que o processo de interpolação afeta a estabilidade do esquema. Consequentemente, é necessário reduzir o intervalo de tempo.

Como mencionado anteriormente, o mC -SSPMS também foi comparado com o método LF2-CPLTS, proposto por [Angulo et al. \(2014\)](#). Para isso, o LF2-CPLTS foi implementado

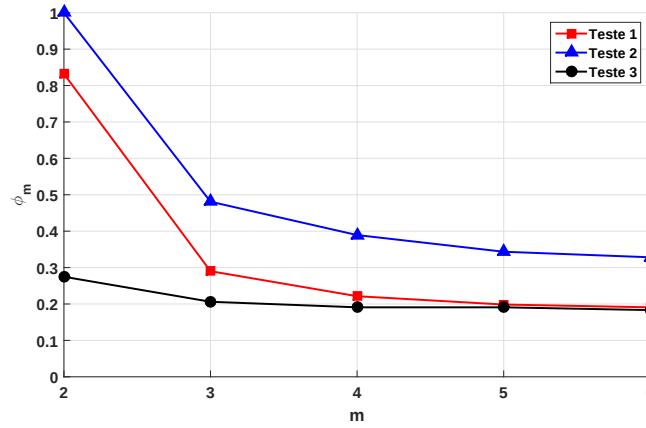


Figura 4.5: Parâmetros de não uniformidade entre classes.

Tabela 4.5: Teste 1: eficiência e acurácia LF2-CPLTS

N	2 camadas		3 camadas		4 camadas		5 camadas	
	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r
1	2,401	0,1650	2,926	0,1659	3,367	0,1547	3,148	0,1670
2	13,114	0,0591	16,309	0,0592	18,509	0,0560	16,133	0,0550
3	25,021	0,0442	31,541	0,0428	29,930	0,0430	31,615	0,0431
4	57,566	0,0261	73,103	0,0237	69,587	0,0243	70,154	0,0263
5	109,963	0,0173	137,799	0,0164	127,600	0,0166	128,768	0,0167

de acordo com a referência e aplicado ao Teste 1, o tempo computacional da simulação foi medido, mantendo a dissipação numérica próxima à apresentada pelos demais métodos (Tabela 4.3). Os resultados são apresentados na Tabela 4.5. Observa-se que o LF2-CPLTS obteve um desempenho muito pior que o mC -SSPMS(6,3), em termos de eficiência computacional. Isso já era esperado já que, apesar do LF2 necessitar de menos operações matemáticas que o SSPMS(6,3), com uso do esquema LTS é preciso que o passo de tempo seja reduzido a valores bem menores para que a dissipação numérica esteja na mesma ordem que o mC -SSPMS(6,3). Esse fato é demonstrado na Tabela 4.6, que traz o fator de redução C_r para o LF2-CPLTS. A fim de facilitar a comparação, os valores apresentados na Tabela 4.6 estão relacionados com o passo de tempo do SSPMS(6,3) padrão.

Como pode ser verificado na Tabela 4.5, o mC -SSPMS foi em média 3 vezes mais rápido que o LF2-CPLTS. Além disso, o LF2-CPLTS se mostrou mais dissipativo, sendo que para controlar essa dissipação o passo de tempo teve de ser reduzido, apresentando em média um coeficiente de redução 2.5 vezes menor que o mC -SSPMS(6,3). Para $N = 5$ o C_r no LF2-CPLTS foi em média 3,4 menor que no mC -SSPMS(6,3).

Nos resultados para o mC -SSPMS(6,3) apresentados até agora, foram utilizadas interpolações lineares para satisfazer o fluxo numérico entre elementos de classes adjacentes. Contudo,

Tabela 4.6: Fator C_r LF2-CPLTS

N	2 camadas	3 camadas	4 camadas	5 camadas
1	0,94	0,46	0,35	0,34
2	0,36	0,19	0,16	0,16
3	0,36	0,17	0,16	0,16
4	0,30	0,17	0,13	0,13
5	0,25	0,15	0,12	0,11

Tabela 4.7: Teste 1: eficiência e acurácia do mC -SSPMS(6,3) com subclasses

N	2C-SSPMS(6,3)		3C-SSPMS(6,3)		4C-SSPMS(6,3)		5C-SSPMS(6,3)	
	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r
1	2.387	1.692	1.942	0.1604	1.707	0.1718	1.802	0.1830
2	6.350	0.0579	5.075	0.0550	4.377	0.0590	4.998	0.0623
3	14.643	0.0442	11.305	0.0416	9.548	0.0434	11.959	0.0451
4	28.403	0.0264	22.166	0.0245	19.278	0.0253	23.688	0.0246
5	47.664	0.0166	37.153	0.0152	30.610	0.0156	40.001	0.0165

essa estratégia pode ser implementada evitando o uso de interpolações. Assim, foi criada uma implementação onde os valores dos campos entre classes adjacentes são calculados, a fim de satisfazer o fluxo. Nessa implementação foram adicionadas subclasses contendo a primeira vizinhança com elementos das classes que fazem fronteira com outras de menor passo de tempo. Os valores nodais dos elementos de cada subclasse são calculados nos mesmos instantes da marcha da classe vizinha. Dessa forma, o fluxo numérico pode ser satisfeito. Essa versão foi aplicada ao teste 1. Os tempos de execução e o erro, para $N = 1, \dots, 5$, são apresentados na Tabela 4.7.

Pode ser observado na Tabela 4.8 que o C_r , em geral, se manteve quase igual ao caso com uso de interpolações. Com uma diferença. Para $m = 2, 3, 4$, C_r teve uma redução, porém essa redução foi a mesma. Isso é importante para que o ganho computacional seja mantido, na medida em que N aumenta. Isso só não ocorreu para $m = 5$. Isso reforça a ideia que um número muito grande de classes afeta a estabilidade do esquema, mesmo sem o uso de interpolações. Como esperado, o tempo de CPU foi um pouco maior. Porém era esperada uma redução no

Tabela 4.8: Teste 1: fator de redução C_r para o mC -SSPMS(6,3) com subclasses

N	2C-SSPMS	3C-SSPMS	4C-SSPMS	5C-SSPMS
1	0,98	0,81	0,81	0,73
2	0,83	0,65	0,65	0,52
3	0,83	0,65	0,65	0,50
4	0,83	0,65	0,65	0,46
5	0,83	0,65	0,65	0,44

erro, o que não ocorreu. Isso mostra que a versão com interpolações foi mais eficiente do ponto de vista de tempo de CPU.

4.2.2 Teste 2: Linha de corrente

O teste anterior abordou um problema onde um pulso gaussiano foi inserido no domínio a partir de condições iniciais do campo elétrico. Muitas vezes é necessário excitar um campo em uma região do domínio. Dessa forma, é importante demonstrar que o esquema proposto também é adequado quando outras formas de excitação de campo são usadas. O campo pode ser excitado de pelo menos três maneiras diferentes: por um esquema de *hard source*, pelo método *total field/scattered field*, ou diretamente pelo fluxo numérico. Nesse teste será simulado uma propagação eletromagnética provocada por uma linha de corrente onde um campo elétrico oscila da forma: $E_z(t) = E_0 \sin(\omega t)$. A imposição do campo é feita usando o fluxo numérico. Para isso basta que $\Delta E = E^- - E_z(t)$. Nesse teste $E_0 = 1 \text{ volt/metro}$ e o campo excitado possui um comprimento de onda $\lambda = 0,5 m$.

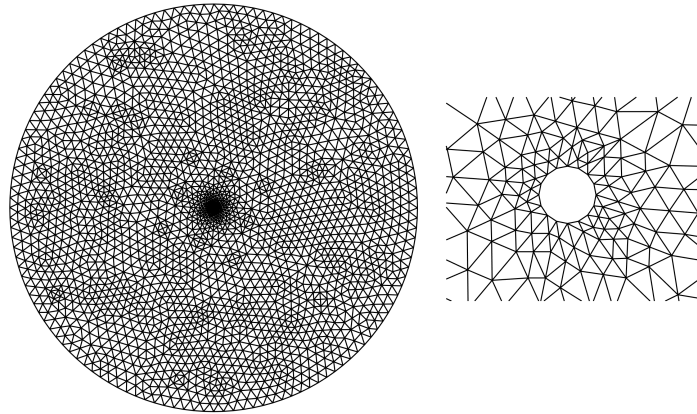


Figura 4.6: Malha teste 2 com 4896 elementos.

Uma malha triangular com 4896 elementos foi criada para representar o domínio computacional, como mostrado na Figura 4.6. O domínio possui um diâmetro de $1 m$. Sendo a fronteira externa circular e o campo irradiado, uma onda plana cilíndrica, foi possível usar a condição Silver Müller de primeira ordem para simular fronteiras abertas sem que houvesse uma reflexão significativa. Como pode ser observado na Figura 4.6, os elementos têm tamanhos muito discrepantes de forma que a razão entre a maior e a menor aresta é 110.25. A solução para o campo elétrico é apresentada na Figura x. O parâmetro ϕ_m calculado com $m = 2, 3, 4, 5, 6$ para a malha, é mostrado na Figura 4.5. Observa-se que ϕ_m tende a saturação de pois de $m = 4$. Portanto espera-se que para $m = 4$ a estratégia multiclasse traga um *speedup* no tempo de CPU. Isso pode ser verificado na Tabela 4.9.

Como pode ser observado na Tabela 4.9, a estratégia 4C-sspmS(6,3) foi em média duas

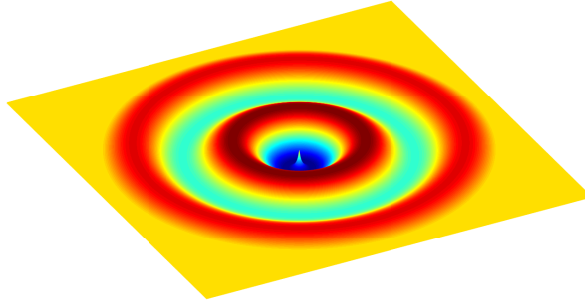
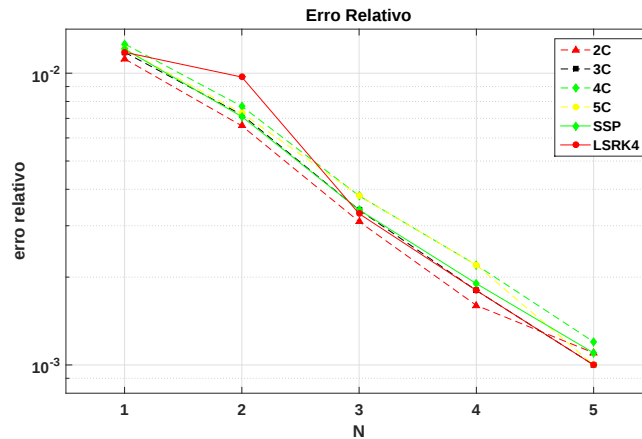


Figura 4.7: Teste 2: solução para o campo elétrico.

Tabela 4.9: Teste 2: eficiência e precisão de mC -SSPMS(6,3), SSPMS(6,3) e RK4

N	2C-SSPMS(6,3)		3C-SSPMS(6,3)		4C-SSPMS(6,3)		5C-SSPMS(6,3)		SSPMS(6,3)		RK4	
	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r	t_{exec}	e_r
1	71.07	0.0112	57.24	0.0118	40.95	0.0126	47.16	0.0121	76.46	0.0121	170.78	0.0118
2	162.62	0.0066	123.20	0.0072	91.34	0.0077	102.43	0.0073	187.03	0.0071	381.22	0.0097
3	312.26	0.0031	239.76	0.0034	183.40	0.0038	203.31	0.0038	362.71	0.0034	801.77	0.0033
4	612.80	0.0016	470.09	0.0018	352.63	0.0022	377.54	0.0022	744.25	0.0019	1475.00	0.0018
5	1134.20	0.0011	834.83	0.0010	641.61	0.0012	723.50	0.0010	1341.60	0.0011	3557.00	0.0010

vezes mais rápido que a versão padrão. Além disso, a estratégia proposta também não alterou a taxa de convergência do esquema DGTD, como pode ser verificado na Tabela 4.9 e na Figura 4.8.

Figura 4.8: Erro relativo *versus* ordem polinomial para o teste 2.

4.2.3 Teste 3: Guia de onda em L baseado em cristal fotônico

No capítulo 2 foram apresentados os conceitos básicos dos cristais fotônicos bidimensionais. Também foram discutidos os conceitos de diagrama de bandas e *gaps* fotônicos. Além

Tabela 4.10: Teste 3: Divisão de elementos entre classes

Elementos por classe	2C-SSPMS(6,3)	3C-SSPMS(6,3)	4C-SSPMS(6,3)
Classe 1	19417	9103	1402
Classe 2	16107	10314	7701
Classe 3		16107	10314
Classe 4			16107

disso, inserindo defeitos lineares, foi demonstrada uma maneira simples de projetar guias de ondas baseados em cristais fotônicos.

Nesse teste, os conceitos abordados no capítulo 2 foram colocados em prática com o objetivo de projetar um guia de ondas com uma curva em L. A base do cristal fotônico usada aqui é composta por pilares dielétricos circulares, $\epsilon_1 = 11,4$, imersas em ar, $\epsilon_2 = 1$, formando uma matriz quadrada com constante de rede $a = 0,57 \mu\text{m}$. O raio dos pilares é $r_a = 0,114 \mu\text{m}$. O cristal fotônico tem um *band gap* fotônico para o campo com polarização TM, cuja faixa de frequência é estendida de $0,35(= \omega a/2\pi c)$ a $0,42$. O guia de onda óptico é criado removendo uma fileira de pilares onde a luz é confinada. O pulso incidente, com banda de frequência dentro do *gap*, é considerado como:

$$Ez(x, y, t) = E_0 \cos\left(\frac{\pi y}{d}\right) \cos(2\pi f_m t) e^{-\left(\frac{t-t_0}{2\sigma}\right)^2} \quad (4.4)$$

onde $d = 2(a - r_a)$ é a largura do guia de onda, $\sigma = 5 \times 10^{-14}$ s define a largura de banda do pulso e $f_m = 2 \times 10^{14}$ Hz é a frequência central. Além disso, nas simulações, foi considerado $t_0 = 2,5 \times 10^{-13}$ s.

A Figura 4.9 (a) mostra a malha triangular que representa o domínio computacional. A condição absorvente de primeira ordem de Silver-Muller é aplicada na fronteira do domínio computacional. Foi utilizada uma malha não uniforme que consiste em 17935 vértices e 35524 triângulos. A razão entre as arestas maior e a menor é de 13,5810 e, neste caso, o intervalo que contém os passos de tempo locais é dado por $(2/3c)r_G[4, 20, 60, 45] \times 10^{-9}$ s. As simulações foram realizadas usando o *mC*-SSPMS(6,3) com $m = 2, 3, 4$ para $N = 1, 2, 3, 4, 5$. A Tabela 4.10 mostra a divisão dos elementos em cada classe. A solução para o componente do campo elétrico, E_z , é apresentada na Figura 4.9 (b). Além disso, a Figura 4.9 (c) é o gráfico *quiver* do vetor Poynting que indica a direção de propagação. Os tempos de CPU são mostrados na Tabela 4.11, o desempenho do SSPMS(6,3) padrão e o RK4 também são relatados.

Mais uma vez, nota-se uma aceleração no tempo computacional. Isto ocorre quando $m = 2, 3$. O ganho G_p , é mostrado na Tabela 4.11. Observa-se que a principal melhoria de eficiência ocorreu por $m = 3$. O 3C-SSPMS foi em média 1,53 vezes mais rápido que a versão

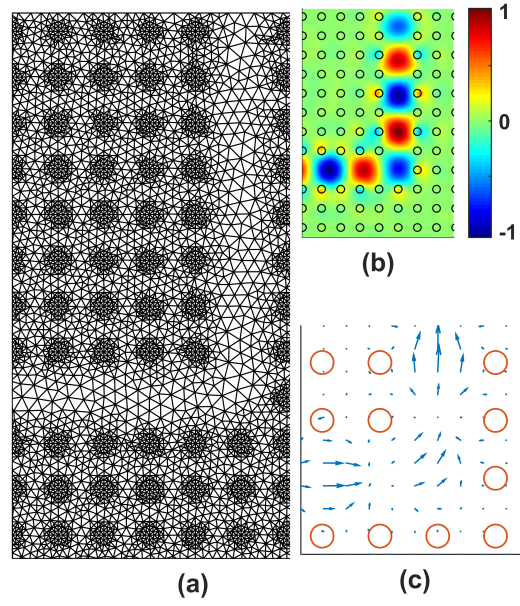


Figura 4.9: (a) Malha multiescala representando um guia de ondas de cristal fotônico. (b) Solução E_z para o 3C-SSPMS(6,3) com $N = 5$. (c) Vetor de Poynting.

Tabela 4.11: Teste 3: eficiência e acurácia para mC -SSPMS(6,3), SSPMS(6,3) and RK4.

N	2C-SSPMS(6,3)			3C-SSPMS(6,3)			4C-SSPMS(6,3)			SSPMS(6,3)	RK4
	t_{exec}	G_p	C_r	t_{exec}	G_p	C_r	t_{exec}	G_p	C_r	t_{exec}	t_{exec}
1	408	1.69	1.00	375	1.84	1.00	447	1.54	0.83	690	1410
2	1553	1.37	0.85	1339	1.58	0.85	1709	1.24	0.65	2125	3942
3	4032	1.30	0.79	3415	1.54	0.77	4482	1.17	0.58	5251	8815
4	8180	1.27	0.75	7695	1.35	0.75	9832	1.06	0.56	10389	18300
5	14470	1.28	0.73	13815	1.34	0.73	18421	1.01	0.54	18570	33264

standard e 2,81 comparado ao método RK4. Enquanto N aumenta, o passo de tempo do método mC -SSPMS tende a se tornar mais restritivo. Tabela 4.11 apresenta os valores experimentais para C_r . É importante enfatizar que o aumento excessivo no número de classes pode degradar a estabilidade do esquema. Esse comportamento foi observado nos experimentos numéricos e é ilustrado pelos valores de C_r .

Observando a malha apresentada na Figura 4.9 (a), pode ser observado que os elementos têm aproximadamente três tamanhos distintos. Portanto, intuitivamente, o 3C-SSPMS parece ser o mais adequado para este problema. A Figura 4.5 apresenta o parâmetro ϕ_m para o Teste 2. Mais uma vez, percebe-se que ϕ_m tende a saturação. O valor $m = 3$ é o primeiro antes do regime de saturação. Isso indica que quando $m > 3$ é possível haver ganho computacional. Isso realmente acontece, como pode ser observado na Tabela 4.11. O parâmetro ϕ_m não determina um número ótimo de classes. No entanto, ϕ_m produz um conhecimento prévio que ajuda a

definir o número de classes para obter um aumento de velocidade.

É de grande interesse conhecer a dispersão que o guia impõe ao sinal propagado. Para isso, o campo elétrico foi aferido colocando um detector na saída do guia de ondas. Uma FFT foi calculada a partir do sinal detectado no domínio do tempo. A Figura 4.10 mostra uma comparação entre o sinal original e o sinal propagado pelo guia de ondas. Nota-se o sinal propagado sofreu dispersão ao longo do guia. Como pode ser verificado, o pico do pulso foi deslocado da frequência central original. A curva mostrada na Figura 4.10 (a) foi obtida com o método 3C-SSPMS(6,3) com ordem polinomial $N = 3$.

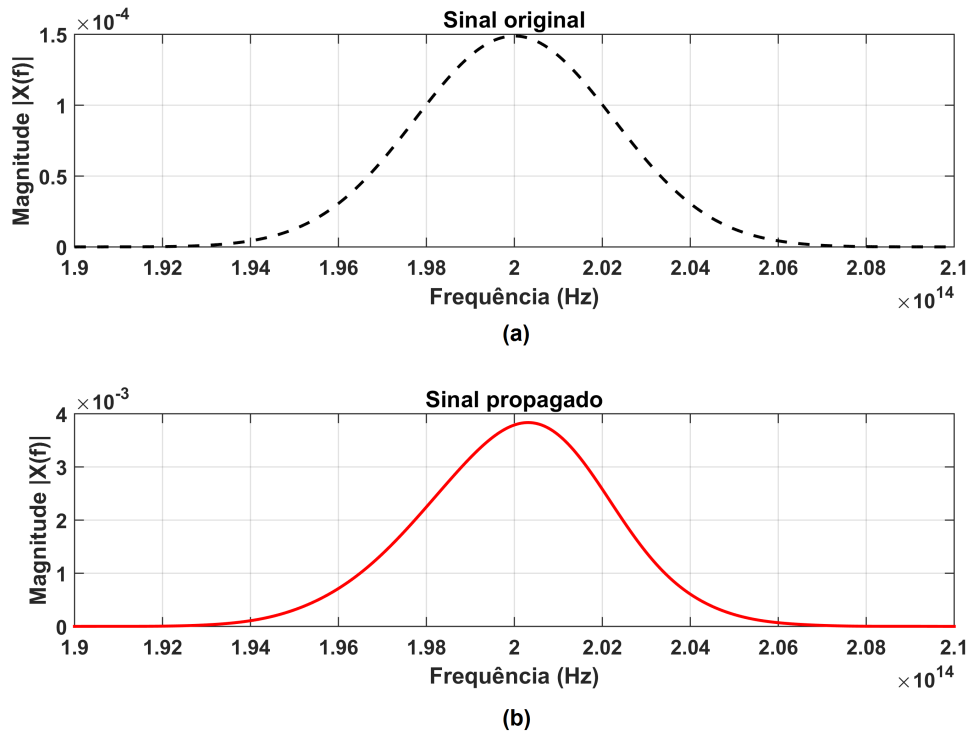


Figura 4.10: Dispersão sofrida pelo pulso ao se propagar pelo guia de ondas

4.2.4 Teste 4: Porta lógica *or*

O dispositivo usado nesse teste é uma porta lógica, *or*, apresentada por [D'souza e Mathew \(2016\)](#). A porta lógica é formada por três guias de ondas acoplados a um anel ressonador. No projeto foi usado como base, um cristal fotônico formado por colunas dielétricas, $\epsilon_r = 11,4$, imersas em ar. A constante de rede é $a = 0,650 \mu\text{m}$, e o raio das colunas, $r = 0,230 \mu\text{m}$. Os guias e o anel são formados retirando fileiras de colunas dielétricas. A Figura 4.11 (a) traz a representação geométrica do dispositivo. A fim de simular fronteiras livres, uma região com PML de largura $10a$ foi incluída. O dispositivo foi projetado para operar em alguns comprimentos

de onda específicos. Nesse teste, o comprimento de onda foi considerado $\lambda_0 = 1,524 \mu\text{m}$. A malha que representa o domínio computacional é apresentada na Figura 4.11 (b). Essa malha possui 25334 elementos. A razão entre a maior e a menor aresta é 125. Os passos de tempo referentes a cada elemento variam entre $(2/3c)r_G[8, 41; 841, 17] \times 10^{-9}$ s.

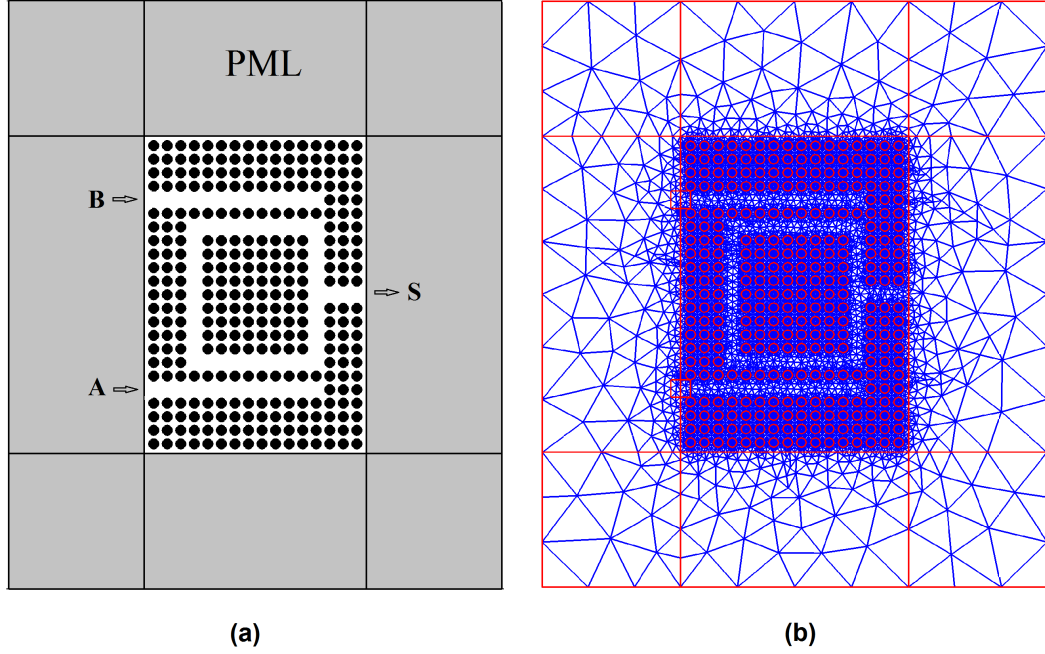


Figura 4.11: (a): representação geométrica da porta *or* e da região de PML que a cerca. (b): malha não estruturada representando o domínio computacional.

Foram simulados os três cenários possíveis. Quando somente a entrada *A* é excitada com um campo eletromagnético, representando $A = 1$ e $B = 0$. Quando somente *B* é excitada, representando a operação lógica $A = 0$ e $B = 1$. Por fim, quando as duas entradas são excitadas, representando $A = 1$ e $B = 1$. Nas três situações, um campo eletromagnético deve existir na saída, representando o nível lógico 1. Isso realmente ocorreu, como pode ser observado na Figura 4.13, que é o resultado da simulação usando o 3C-SSPMS(6,3) com $N = 4$. A janela de tempo utilizada foi de $50 \mu\text{s}$.

Nesse teste, o parâmetro ϕ_m desenvolvido anteriormente foi usado para determinar um número de classes que possa trazer um ganho de tempo de CPU. Para isso, foi utilizado um esquema de divisão de classes com valores fixos para cada Δt_k . Recursivamente, a partir de 2, m foi incrementado e o parâmetro ϕ_m foi calculado. Foi estabelecido um valor de *threshold*, $\tau < 0,1$, para interromper a evolução de m e é assumido que o valor adequado é $m - 1$. Aqui, o número de classes encontrado com esse procedimento foi $m = 3$. Os valores de ϕ_m normalizados são mostrados na Figura 4.13.

O problema foi simulado usando o esquema 3C-SSPMS(6,3), além do método convencional SSPMS(6,3). Todas as simulações foram realizadas com $N = 4$. A Tabela 4.12 traz o tempo gasto para cada esquema. O 3C-SSPMS(6,3) trouxe uma economia de tempo de 24%

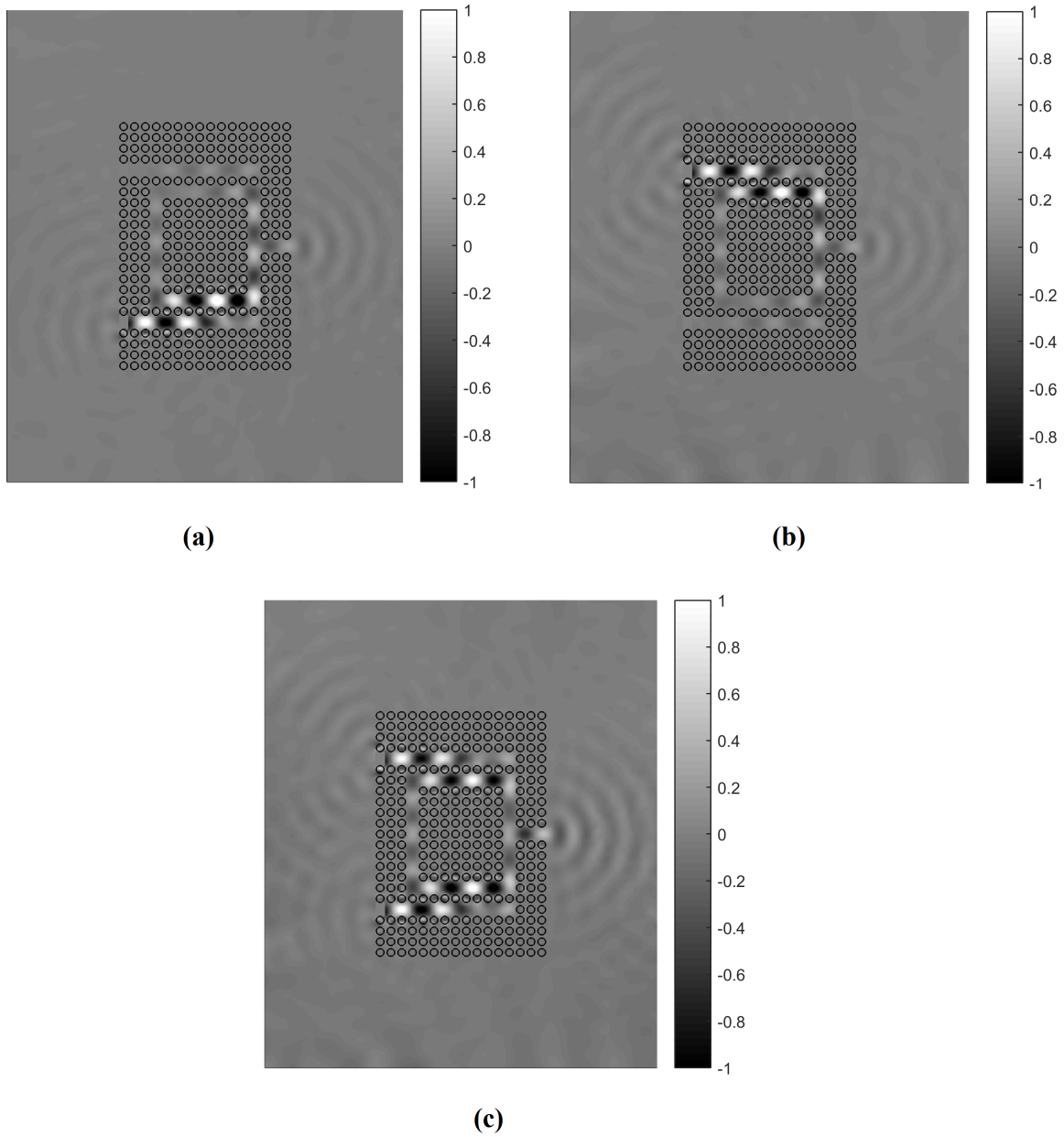


Figura 4.12: Campo elétrico, E_z , para teste 3 usando o esquema 3C-SSPMS(6,3) com $N = 4$. A janela de tempo utilizada foi de $50 \mu s$. (a) solução para $A = 1, B = 0$. (b) solução para $A = 0$ e $B = 1$. (c) solução para $A = 1$ e $B = 1$.

em relação ao método convencional. É importante ressaltar que, como visto na Tabela 4.12, o número de elementos na classe de menor passo de tempo é muito maior que nas outras classes, e mesmo assim o esquema 3C-SSPMS trouxe ganho computacional. Apesar do ganho parecer modesto, em simulações mais longas, envolvendo por exemplo dispositivos em cascata, uma economia de tempo de 24% pode se tornar muito significativa. Um fato a ser observado é que 75% dos elementos foram classificados na classe com menor passo de tempo. Nos testes 1 e 2, esse número foi de 5% e 45%. Isso reforça o fato de que a distribuição dos elementos nas classes é um dos fatores determinantes para se obter ganho computacional.

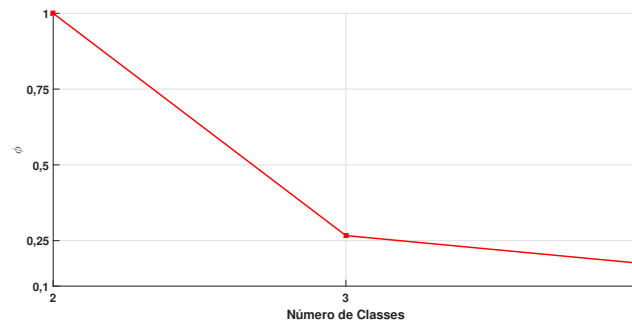


Figura 4.13: Parâmetro de não uniformidade entre classes.

Tabela 4.12: Teste 4: Resultados da simulação dos esquemas *mC*-SSPMS(6,3), SSPMS(6,3)

	3C-SSPMS(6,3)	SSPMS(6,3)
tempo (horas)	3.1272	4.1092
n° elementos classe 1	2092	
n° elementos classe 2	4272	
n° elementos classe 3	18970	

A fim de estimar o tempo de resposta da porta lógica, foi medida a potência em função do tempo na porta S , quando $A = 1$ e $B = 0$. A potência instantânea no intervalo de $4,1 \times 10^{-13}$ s a $4,13 \times 10^{-13}$ s é mostrada na Figura 4.14. Esse intervalo de tempo equivale a aproximadamente meio período da senoide. A potência média medida nesse intervalo foi de -32.16 dBm. Essa potência, apesar de baixa, pode ser detectada por um fotodiodo PIN com sensibilidade na faixa de -40 a -30 dBm (Wentworth (2009)). Para o caso que $A = 0$ e $B = 1$ espera-se que a potência média seja muito próxima devido a simetria da porta. Quando $A = 1$ e $B = 1$ é esperado uma potência superior para o mesmo intervalo de tempo. Portanto, o tempo mínimo de resposta dessa porta lógica, quando usado um fotodiodo PIN como detector, é de $4,13 \times 10^{-13}$ s.

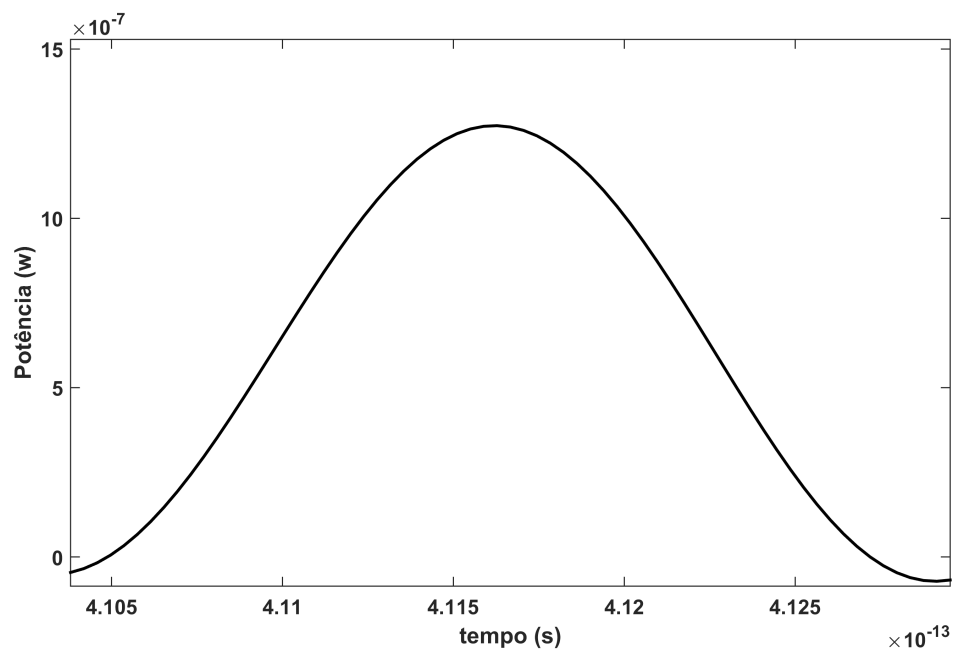


Figura 4.14: Potência instantânea

Conclusão e Trabalhos Futuros

5.1 Conclusão

Na presente tese foram desenvolvidas modelagens matemáticas, baseadas no método de Galerkin, para problemas relacionados à cristais fotônicos bidimensionais. No tratamento das estruturas de bandas, foi desenvolvida uma maneira simples e eficiente de impor as condições periódicas inerentes ao teorema de Bloch. Também foi desenvolvida uma estratégia de passo local de tempo baseada no método SSPMS, denominada *mC-SSPMS*. A estratégia proposta tem o objetivo de aumentar a eficiência computacional em simulações de dispositivos baseados em cristal fotônicos, usando o DGTD.

No levantamento do diagrama de estrutura de bandas de um cristal fotônico, foi resolvido um problema de autovalor onde as frequências são determinadas. Nesse trabalho, foi aplicada uma modelagem variacional, baseada no método de Galerkin. Devido ao teorema de Bloch, tanto a função de forma, quanto a função de teste, devem ser periódicas em relação a constante de rede do cristal. Para evitar o uso de funções periódicas, que são mais complicadas que os polinômios geralmente usadas no FEM, foi proposta uma forma de impor a periodicidade da função de teste.

A estratégia consiste em duplicar a célula unitária, de forma que, os graus de liberdade na segunda tenham sua coordenada x como uma translação $x + a$ em relação aos graus de liberdade na célula original. Em seguida foi aplicada uma fatoração no sistema resultante, diminuindo o número de graus de liberdade a menos que antes da duplicação. O sistema se conserva Hermitiano. A eficiência da estratégia é confirmada com exemplos comparados com a literatura.

Uma outra contribuição dessa tese é o desenvolvimento do esquema *mC-SSPMS*. Essa estratégia multi classe, baseada no método multi passo linear de estabilidade forte, foi aplicada à discretização DG formando um método DGTD eficiente computacionalmente quando aplicado à problemas com múltiplas escalas. A fim de avaliar a proposta, problemas de propagação eletromagnética foram simulados. Os problemas simulados foram: a propagação de um campo gerado por uma linha de corrente, a propagação de um pulso gaussiano em um guia com curva de 90° e uma porta lógica, *or*, baseada em cristal fotônico. Nos resultados, o esquema proposto foi comparado com sua versão padrão, além do RK4 e o método LF2-CPLTS.

No problema do pulso gaussiano, sem perda de precisão, o 4C-SSPMS foi em média 2,5 vezes mais rápido que sua versão padrão. Em relação ao RK4, o 4C-SSPMS foi 4,24 vezes mais rápido. Na comparação com o método LF2-CPLTS, o 4C-SSPMS foi em média 3 vezes mais rápido. Os resultados no teste 2 foram similares aos resultados do teste 1. No terceiro teste, a versão 3C-SSPMS foi em média 1,5 vezes mais rápido que a versão padrão SSPMS e 2,8 vezes mais rápido que o RK4. Na simulação da porta *or*, o 3C-SSPMS proporcionou uma economia de 24% em relação a versão SSPMS padrão.

Nesse trabalho também foi proposta uma forma de ter um conhecimento prévio sobre quantas classes devem ser usadas a fim de promover uma melhoria na eficiência computacional. Para isso, foi proposto um parâmetro de malha baseado no passo de tempo das classes. Esse parâmetro é dependente das características de cada malha e tende à saturação com o aumento do número de classes. Foi observado que quando há a saturação, o aumento do número de classes não traz melhoria na eficiência computacional.

5.2 Trabalhos Futuros

A estratégia proposta aqui mostrou que pode trazer ganho de tempo computacional quando aplicada a problemas com múltiplas escalas. Contudo, ainda há muito a explorar. Portanto, como possíveis trabalhos futuros, pode-se apontar:

- Criar uma versão da estratégia muticlasse onde a ordem polinomial, da discretização espacial de cada classe, possa ser distinta. Dessa forma, seria possível usar uma discretização espacial com ordem variável ou até adaptativa. Isso poderia aumentar o ganho de tempo computacional.
- Criar uma versão DGTD com integração do tempo adaptativa.
- Encontrar formas de otimizar a classificação dos elementos. É possível que as características da malha possam ser usadas para determinar uma classificação ótima para os elementos.
- Fazer uma análise de estabilidade e determinar uma expressão que defina o passo de tempo máximo do esquema. Além do interesse em entender de forma mais completa as características de estabilidade, encontrar uma expressão para o passo de tempo é importante para que a aplicação da estratégia seja mais direta.
- Tentar incluir limitadores na formulação a fim de se obter a propriedade SSP do método linear multipasso com LTS. Isso permitiria a aplicação da estratégia em problemas cuja soluções possuam descontinuidades.

Referências Bibliográficas

- Alvarez et al.(2015)** Jesus Alvarez, Luis Diaz Angulo, Amelia Rubio Bretones, Carlos M de Jong van Coevorden, e Salvatore G Garcia. Efficient antenna modeling by dgtd: Leap-frog discontinuous galerkin timedomain method. *IEEE Antennas and Propagation Magazine*, 57 (3):95–106. Citado na pág.
- Álvarez González(2014)** Jesús Álvarez González. *A discontinuous Galerkin finite element method for the Time-Domain solution of Maxwell equations*. Universidad de Granada. Citado na pág.
- Angulo et al.(2014)** LD Angulo, J Alvarez, FL Teixeira, MF Pantoja, e SG Garcia. Causal-path local time-stepping in the discontinuous galerkin method for maxwell’s equations. *Journal of Computational Physics*, 256:678–695. Citado na pág.
- Arfken e Weber(1999)** George B Arfken e Hans J Weber. Mathematical methods for physicists, 1999. Citado na pág.
- Axmann e Kuchment(1999)** Waldemar Axmann e Peter Kuchment. An efficient finite element method for computing spectra of photonic and acoustic band-gap materials: I. scalar case. *Journal of Computational Physics*, 150(2):468–481. Citado na pág.
- Baba(2007)** Toshihiko Baba. Remember the light. *Nature Photonics*, 1:11–12. Citado na pág.
- Baba(2008)** Toshihiko Baba. Slow light in photonic crystals. *Nature photonics*, 2(8):465–473. Citado na pág.
- Booth(1954)** Andrew D Booth. Bashforth–adams method for the numerical solution of differential equations. *Nature*, 173(4405):635. Citado na pág.
- Bresten et al.(2017)** Christopher Bresten, Sigal Gottlieb, Zachary Grant, Daniel Higgs, David Ketcheson, e Adrian Németh. Explicit strong stability preserving multistep runge–kutta methods. *Mathematics of Computation*, 86(304):747–769. Citado na pág.
- Bruyère et al.(2008)** A Bruyère, L Illoul, Francisco Chinesta, e S Clenet. Comparison between nem and fem in 2-d magnetostatics using an error estimator. *IEEE Transactions on Magnetics*, 44(6):1342–1345. Citado na pág.

- Burger et al.(2005)** Sven Burger, Roland Klose, Achim Schaedle, Frank Schmidt, e Lin W Zschiedrich. Fem modeling of 3d photonic crystals and photonic crystal waveguides. Em *Integrated Optoelectronic Devices 2005*, páginas 164–173. International Society for Optics and Photonics. Citado na pág.
- Busch et al.(2011)** Kurt Busch, Michael König, e Jens Niegemann. Discontinuous galerkin methods in nanophotonics. *Laser & Photonics Reviews*, 5(6):773–809. Citado na pág.
- Chen e Liu(2013)** Jiefu Chen e Qing Huo Liu. Discontinuous galerkin time-domain methods for multiscale electromagnetic simulations: A review. *Proceedings of the IEEE*, 101(2):242–254. Citado na pág.
- Chen e Babuška(1995)** Qi Chen e Ivo Babuška. Approximate optimal points for polynomial interpolation of real functions in an interval and in a triangle. *Computer Methods in Applied Mechanics and Engineering*, 128(3-4):405–417. Citado na pág.
- Chihara(2011)** Theodore S Chihara. *An introduction to orthogonal polynomials*. Courier Corporation. Citado na pág.
- Cohen et al.(2006)** Gary Cohen, Xavier Ferrieres, e Sébastien Pernet. A spatial high-order hexahedral discontinuous galerkin method to solve maxwell's equations in time domain. *Journal of Computational Physics*, 217(2):340–363. Citado na pág.
- Colman et al.(2010)** Pierre Colman, Chad Husko, Sylvain Combrié, Isabelle Sagnes, Chee Wei Wong, e Alfredo De Rossi. Temporal solitons and pulse compression in photonic crystal waveguides. *Nature Photonics*, 4(12):862–868. Citado na pág.
- Contu et al.(2012)** Pietro Contu, Cornelis Van Der Mee, e Sebastiano Seatzu. A finite element frequency domain method for 2d photonic crystals. *Journal of Computational and Applied Mathematics*, 236(16):3956–3966. Citado na pág.
- Diaz e Grote(2009)** Julien Diaz e Marcus J Grote. Energy conserving explicit local time stepping for second-order wave equations. *SIAM Journal on Scientific Computing*, 31(3):1985–2014. Citado na pág.
- Díaz Angulo(2015)** Luis Manuel Díaz Angulo. Time domain discontinuous galerkin methods for maxwell equations. Citado na pág.
- Dokainish e Subbaraj(1989)** MA Dokainish e K Subbaraj. A survey of direct time-integration methods in computational structural dynamics—i. explicit methods. *Computers & Structures*, 32(6):1371–1386. Citado na pág.
- D'souza e Mathew(2016)** Nirmala Maria D'souza e Vincent Mathew. Interference based square lattice photonic crystal logic gates working with different wavelengths. *Optics & Laser Technology*, 80:214–219. Citado na pág.

- Dulkeith et al.(2005)** Eric Dulkeith, Sharee J McNab, e Yurii A Vlasov. Mapping the optical properties of slab-type two-dimensional photonic crystal waveguides. *Physical Review B*, 72 (11):115102. Citado na pág.
- Fahs(2009)** Hassan Fahs. High-order leap-frog based discontinuous galerkin method for the time-domain maxwell equations on non-conforming simplicial meshes. *Numerical mathematics*, 2(3):275–300. Citado na pág.
- Fu et al.(2013)** Yulan Fu, Xiaoyong Hu, e Qihuang Gong. Silicon photonic crystal all-optical logic gates. *Physics letters A*, 377(3-4):329–333. Citado na pág.
- Gallinet et al.(2015)** Benjamin Gallinet, Jérémy Butet, e Olivier JF Martin. Numerical methods for nanophotonics: standard problems and future challenges. *Laser & Photonics Reviews*, 9 (6):577–603. Citado na pág.
- Gottlieb e Shu(1998)** Sigal Gottlieb e Chi-Wang Shu. Total variation diminishing runge-kutta schemes. *Mathematics of computation of the American Mathematical Society*, 67(221):73–85. Citado na pág.
- Gottlieb et al.(2011)** Sigal Gottlieb, David I Ketcheson, e Chi-Wang Shu. *Strong stability preserving Runge-Kutta and multistep time discretizations*. World Scientific. Citado na pág.
- Harten(1983)** Ami Harten. High resolution schemes for hyperbolic conservation laws. *Journal of computational physics*, 49(3):357–393. Citado na pág.
- Hesthaven(1998)** Jan S Hesthaven. From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex. *SIAM Journal on Numerical Analysis*, 35(2):655–676. Citado na pág.
- Hesthaven e Warburton(2002a)** Jan S Hesthaven e Tim Warburton. Nodal high-order methods on unstructured grids: I. time-domain solution of maxwell’s equations. *Journal of Computational Physics*, 181(1):186–221. Citado na pág.
- Hesthaven e Warburton(2007)** Jan S Hesthaven e Tim Warburton. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*. Springer Science & Business Media. Citado na pág.
- Hesthaven e Warburton(2002b)** Jan S Hesthaven e Timothy Warburton. Nodal high-order methods on unstructured grids: I. time-domain solution of maxwell’s equations. *Journal of Computational Physics*, 181(1):186–221. Citado na pág.
- Hirsch(2007)** Charles Hirsch. *Numerical computation of internal and external flows: The fundamentals of computational fluid dynamics*. Elsevier. Citado na pág.

- Joannopoulos et al.(2011)** John D Joannopoulos, Steven G Johnson, Joshua N Winn, e Robert D Meade. *Photonic crystals: molding the flow of light*. Princeton university press. Citado na pág.
- Kahnert(2003)** F Michael Kahnert. Numerical methods in electromagnetic scattering theory. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 79:775–824. Citado na pág.
- Kappeler et al.(2011)** R Kappeler, P Kaspar, e H Jäckel. Propagation loss computation of w1 photonic crystal waveguides using the cutback technique with the 3d-fdtd method. *Photonics and Nanostructures-Fundamentals and Applications*, 9(3):235–247. Citado na pág.
- Karle et al.(2004)** TJ Karle, YJ Chai, CN Morgan, IH White, e TF Krauss. Observation of pulse compression in photonic crystal coupled cavity waveguides. *Journal of lightwave technology*, 22(2):514–519. Citado na pág.
- Ketcheson(2009)** David I Ketcheson. *High order strong stability preserving time integrators and numerical wave propagation methods for hyperbolic PDEs*. University of Washington. Citado na pág.
- Kittel et al.(1976)** Charles Kittel et al. *Introduction to solid state physics*, volume 8. Wiley New York. Citado na pág.
- Koshiba(2001)** Masanori Koshiba. Wavelength division multiplexing and demultiplexing with photonic crystal waveguide couplers. *journal of lightwave technology*, 19(12):1970. Citado na pág.
- Lesaint e Raviart(1974)** Pierre Lesaint e Pierre-Arnaud Raviart. On a finite element method for solving the neutron transport equation. *Mathematical aspects of finite elements in partial differential equations*, (33):89–123. Citado na pág.
- Mandelstam e Taylor(1997)** Vladimir A Mandelshtam e Howard S Taylor. Harmonic inversion of time signals and its applications. *The Journal of chemical physics*, 107(17):6756–6769. Citado na pág.
- Montseny et al.(2008)** Emmanuel Montseny, Sébastien Pernet, Xavier Ferrières, e Gary Cohen. Dissipative terms and local time-stepping improvements in a spatial high order discontinuous galerkin scheme for the time-domain maxwell's equations. *Journal of Computational Physics*, 227(14):6795–6820. Citado na pág.
- Mori e Baba(2004)** Daisuke Mori e Toshihiko Baba. Dispersion-controlled optical group delay device by chirped photonic crystal waveguides. *Applied physics letters*, 85(7):1101–1103. Citado na pág.

- Nicomedes et al.(2012)** Williams Nicomedes, Renato Mesquita, e Fernando Moreira. Calculating the band structure of photonic crystals through the meshless local petrov-galerkin (mlpg) method and periodic shape functions. *IEEE Transactions on Magnetics*, 48(2):551–554. Citado na pág.
- Painter et al.(1999)** Oskar Painter, RK Lee, Axel Scherer, A Yariv, JD O'brien, PD Dapkus, e I Kim. Two-dimensional photonic band-gap defect mode laser. *Science*, 284(5421):1819–1821. Citado na pág.
- Piperno(2006)** Serge Piperno. Symplectic local time-stepping in non-dissipative dgtd methods applied to wave propagation problems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 40(5):815–841. Citado na pág.
- Polycarpou(2005)** Anastasis C Polycarpou. Introduction to the finite element method in electromagnetics. *Synthesis Lectures on Computational Electromagnetics*, 1(1):1–126. Citado na pág.
- Rao(1999)** Sadasiva M Rao. *Time domain electromagnetics*. Academic Press. Citado na pág.
- Reed e Hill(1973)** Wm H Reed e TR Hill. Triangularmesh methodsfor the neutrontransportequation. *Los Alamos Report LA-UR-73-479*. Citado na pág.
- Ruuth e Hundsdorfer(2005)** Steven J Ruuth e Willem Hundsdorfer. High-order linear multistep methods with general monotonicity and boundedness properties. *Journal of Computational Physics*, 209(1):226–248. Citado na pág.
- Sármány et al.(2007)** D Sármány, Mike A Botchev, e Jaap JW van der Vegt. Dispersion and dissipation error in high-order runge-kutta discontinuous galerkin discretisations of the maxwell equations. *Journal of Scientific Computing*, 33(1):47–74. Citado na pág.
- Schmidt e Kauf(2009)** Kersten Schmidt e Peter Kauf. Computation of the band structure of two-dimensional photonic crystals with hp finite elements. *Computer methods in applied mechanics and engineering*, 198(13-14):1249–1259. Citado na pág.
- Sibilia et al.(2008)** Concita Sibilia, Trevor M Benson, Marian Marciniak, e Tomasz Szoplik. *Photonic crystals: physics and technology*. Springer. Citado na pág.
- Sibson(1980)** Robin Sibson. A vector identity for the dirichlet tessellation. Em *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 87, páginas 151–155. Cambridge Univ Press. Citado na pág.
- Sinha et al.(2011)** Ravindra K Sinha, R Bhattacharyya, et al. Photonic crystal slab waveguide-based infiltrated liquid sensors: design and analysis. *Journal of Nanophotonics*, 5(1):053505–053505. Citado na pág.

- Skivesen et al.(2007)** Nina Skivesen, Amølie Têtu, Martin Kristensen, Jørgen Kjems, Lars H Frandsen, e Peter I Borel. Photonic-crystal waveguide biosensor. *Optics Express*, 15(6): 3169–3176. Citado na pág.
- Soljačić et al.(2002)** Marin Soljačić, Steven G Johnson, Shanhui Fan, Mihai Ibanescu, Erich Ippen, e JD Joannopoulos. Photonic-crystal slow-light enhancement of nonlinear phase sensitivity. *JOSA B*, 19(9):2052–2059. Citado na pág.
- Sukumar et al.(2001)** N Sukumar, B Moran, A Yu Semenov, e VV Belikov. Natural neighbour galerkin methods. *International journal for numerical methods in engineering*, 50(1):1–27. Citado na pág.
- Sukumar(1998a)** Natarajan Sukumar. *The natural element method in solid mechanics*. Tese de Doutorado, Northwestern University. Citado na pág.
- Sukumar(1998b)** Natarajan Sukumar. *The natural element method in solid mechanics*. Tese de Doutorado, Northwestern University. Citado na pág.
- Taflove e Hagness(2005)** Allen Taflove e Susan C Hagness. *Computational electrodynamics: the finite-difference time-domain method*. Artech house. Citado na pág.
- Taflove et al.(1995)** Allen Taflove, Susan C Hagness, et al. *Computational electrodynamics: the finite-difference time-domain method*. Norwood, 2nd Edition, MA: Artech House, 1995. Citado na pág.
- Taylor et al.(2000)** Mark A Taylor, Beth A Wingate, e Rachel E Vincent. An algorithm for computing fecke points in the triangle. *SIAM Journal on Numerical Analysis*, 38(5):1707–1720. Citado na pág.
- Villeneuve et al.(1998)** PR Villeneuve, S Fan, SG Johnson, e JD Joannopoulos. Three-dimensional photon confinement in photonic crystals of low-dimensional periodicity. *IEE Proceedings-Optoelectronics*, 145(6):384–390. Citado na pág.
- Watson(1981)** David F Watson. Computing the n-dimensional delaunay tessellation with application to voronoi polytopes. *The computer journal*, 24(2):167–172. Citado na pág.
- Wentworth(2009)** Stuart M Wentworth. *Eletromagnetismo aplicado: abordagem antecipada das linhas de transmissão*. Bookman Editora. Citado na pág.
- Williamson(1980)** JH Williamson. Low-storage runge-kutta schemes. *Journal of Computational Physics*, 35(1):48–56. Citado na pág.
- Yang et al.(2013)** He Yang, Fengyan Li, e Jianxian Qiu. Dispersion and dissipation errors of two fully discrete discontinuous galerkin methods. *Journal of Scientific Computing*, 55(3): 552–574. Citado na pág.

Yee et al.(1966) Kane S Yee *et al.* Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. *IEEE Trans. Antennas Propag*, 14(3): 302–307. Citado na pág.