

Alexandre Ashade Lassance Cunha

Método de solução analítica de equações de autovalores de operadores diferenciais por programação genética, com aplicação na análise de propagação eletromagnética em colunas de produção de óleo parametrizada pelo raio e o percentual de incrustações

TESE DE DOUTORADO

DEPARTAMENTO DE ENGENHARIA ELÉTRICA
Programa de Pós-graduação em Engenharia
Elétrica

Rio de Janeiro
Setembro de 2017



Alexandre Ashade Lassance Cunha

Método de solução analítica de equações de autovalores de operadores diferenciais por programação genética, com aplicação na análise de propagação eletromagnética em colunas de produção de óleo parametrizada pelo raio e o percentual de incrustações

Tese de Doutorado

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-graduação em Engenharia Elétrica da PUC-Rio.

Orientador: Prof. Marco Aurélio Cavalcanti Pacheco

Rio de Janeiro
Setembro de 2017



Alexandre Ashade Lassance Cunha

Método de solução analítica de equações de autovalores de operadores diferenciais por programação genética, com aplicação na análise de propagação eletromagnética em colunas de produção de óleo parametrizada pelo raio e o percentual de incrustações

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-graduação em Engenharia Elétrica da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Marco Aurélio Cavalcanti Pacheco

Orientador

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Marco Antonio Guimaraes Dias

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Valmir Carneiro Barbosa

Universidade Federal do Rio de Janeiro – UFRJ

Prof. Antonio Carneiro de Mesquita Filho

Universidade Federal do Rio de Janeiro – UFRJ

Dr. Douglas Mota Dias

Universidade Estadual do Rio de Janeiro – UERJ

Prof. Márcio da Silveira Carvalho

Coordenador Setorial do Centro Técnico Científico – PUC-Rio

Rio de Janeiro, 13 de Setembro de 2017

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Alexandre Ashade Lassance Cunha

Profissional com sólida formação e experiência tanto em pesquisa quanto em desenvolvimento. Especialista em métodos de apoio à decisão, incluindo modelagem preditiva e algoritmos de otimização. Doutorado e mestrado em Engenharia Elétrica, uma patente depositada e com artigos publicados. Experiência como pesquisador do *IBM RESEARCH*. Atuou em projetos de P&D para grandes empresas, como PETROBRAS, LIGHT e TAG. Experiência como professor do CCE-PUC-RJ em cursos de extensão na área de apoio à decisão e óleo & gás. Domina metodologias de data mining, classificação de padrões, previsão, otimização e computação de alto-desempenho utilizando MPI, CUDA e OpenCL. Experiência de 6 anos em desenvolvimento de sistemas de software, utilizando as linguagens C#, C++, Java, Python e a ferramenta MATLAB.

Ficha Catalográfica

Cunha, Alexandre Ashade Lassance

Método de solução analítica de equações de autovalores de operadores diferenciais por programação genética, com aplicação na análise de propagação eletromagnética em colunas de produção de óleo parametrizada pelo raio e o percentual de incrustações / Alexandre Ashade Lassance Cunha; orientador: Marco Aurélio Cavalcanti Pacheco. – Rio de Janeiro: PUC–Rio, Departamento de Engenharia Elétrica, 2017.

v., 135 f: il. ; 30 cm

Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica.

Inclui bibliografia

1. Engenharia Elétrica – Teses. 2. Operadores Diferenciais Lineares;. 3. Autovalores;. 4. Telemetria de poços de petróleo;. 5. Eletromagnetismo computacional;. 6. Solução Analítica Paramétrica. I. Pacheco, Marco Aurélio Cavalcanti. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

Agradecimentos

Agradeço ao meu orientador, por ter investido em minha pesquisa nos últimos anos. Agradeço, também, à minha esposa e família, pelo apoio dado nos últimos anos. Agradeço também ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) por ter financiado por meio de bolsa os meus 4 anos de doutorado.

Resumo

Cunha, Alexandre Ashade Lassance; Pacheco, Marco Aurélio Cavalcanti. **Método de solução analítica de equações de autovalores de operadores diferenciais por programação genética, com aplicação na análise de propagação eletromagnética em colunas de produção de óleo parametrizada pelo raio e o percentual de incrustações.** Rio de Janeiro, 2017. 135p. Tese de Doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Este trabalho apresenta uma abordagem inovadora para calcular autopares de operadores diferenciais (OD), utilizando programação genética (PG) e computação simbólica. Na literatura atual, o Método dos Elementos Finitos (MEF) e o Método das Diferenças Finitas (MDF) são os mais utilizados. Tais métodos usam discretização para converter o OD em uma matriz finita e, por isso, apresentam limitações como perda de acurácia e presença de soluções espúrias. Além disso, se o domínio do OD fosse alterado, os autopares precisariam ser calculados novamente, pois a representação matricial do operador depende dos parâmetros do problema. Nesse contexto, este trabalho propõe evoluir autofunções analiticamente usando PG, sem discretização do domínio. Com isso, é possível incorporar parâmetros, o que torna a solução obtida válida para uma classe de problemas. Este texto descreve o modelo para OD normais, aplicando conceitos de indivíduos multi-árvore e diferenciação simbólica. O modelo evolui auto-funções e, a partir delas, calcula os autovalores empregando a *razão de Rayleigh*. Experimentos baseados em aplicações da Física mostram que a técnica proposta é capaz de encontrar as autofunções analíticas com a acurácia igual ou melhor que as técnicas numéricas supracitadas. Finalmente, a técnica proposta é aplicada ao problema de propagação de ondas eletromagnéticas em poços de petróleo em ULF e UHF. As soluções analíticas são dadas em função do diâmetro e do percentual de incrustações no poço. Os resultados mostram que, para um conjunto suficientemente grande de valores distintos dos parâmetros, a técnica apresenta tempo de execução inferior às técnicas clássicas, mantendo a acurácia destas.

Palavras-chave

Operadores Diferenciais Lineares; Autovalores; Telemetria de poços de petróleo; Eletromagnetismo computacional; Solução Analítica Paramétrica

Abstract

Cunha, Alexandre Ashade Lassance; Pacheco, Marco Aurélio Cavalcanti (Advisor). **Analytical solution of eigenvalue equations by genetic programming, with application in the analysis of electromagnetic propagation in production pipes of oil, parameterized by the radius and the percentage of incrustations**. Rio de Janeiro, 2017. 135p. Tese de doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

This work presents an innovative approach to calculate the eigenpairs of linear differential operators (LDO), employing genetic programming (GP) and symbolic computation. In the current literature, the Finite Element Method (FEM) and the Finite Difference Method (FDM) are more commonly used. Such methods use discretization to convert the LDO to a finite matrix, therefore causing loss of accuracy and presence of spurious solutions. Additionally, if the domain of the LDO was changed, the eigenpairs would need to be recalculated, since the matrix representation of the LDO depends on the parameters of the problem. In this context, this work proposes to evolve eigenfunctions analytically using GP, without domain discretization. Hence, it is possible to incorporate the parameter, which makes a obtained solution valid for a class of problems. This text describes the model for normal LDO, applying concepts of multi-tree individuals and symbolic differentiation. The presented model evolves eigenfunctions and, then, calculates the eigenvalues using the Rayleigh quotient. Experiments based on Physics problems show that the proposed technique is able to find the analytical eigenfunctions with the same accuracy of the numerical techniques mentioned above. Finally, the proposed technique is applied to the problem of propagation of electromagnetic waves in oil wells in ULF and UHF. The analytical solutions are given as a function of the diameter and percentage of CaCO_3 in the well. The results show that, for a sufficiently large set of distinct values of the parameters, the technique presents execution time inferior to the FEM, while maintaining its accuracy.

Keywords

Linear Differential Operators; Eigenvalues; Oil well telemetry; Computational Electromagnetism; Parametric and analytical solutions

Sumário

1	Introdução	16
2	Descrição do Problema e Revisão da Literatura	20
2.1	Formulação Matemática de Problemas de Autovalores	20
2.1.1	Pré-requisitos Matemáticos	21
2.1.1.1	Conjuntos	21
2.1.1.2	Espaço Métrico	22
2.1.1.3	Espaço Linear	22
2.1.1.4	Normas e Produtos Internos	24
2.1.1.5	Espaço de Hilbert	25
2.1.1.6	Ortogonalidade	26
2.1.1.7	Operadores Lineares	26
2.1.1.8	Operadores Limitados	27
2.1.1.9	Adjuntos de Operadores Lineares	27
2.1.1.10	Operadores Positivos e Negativos	28
2.1.2	Problemas de Autovalores de Operadores Diferenciais Lineares	28
2.1.2.1	Considerações Preliminares	28
2.1.2.2	Razão de Rayleigh	30
2.2	Técnicas para Solução de Problemas de Autovalores na Literatura	30
2.2.1	Considerações Iniciais	31
2.2.2	Método das Diferenças Finitas	32
2.2.3	Método dos Elementos Finitos	36
2.3	Emprego de Programação Genética (PG) na Literatura	39
2.4	Contribuições Originais	41
3	Modelo de Solução de Problemas de Autovalores	44
3.1	Introdução à PG Aplicada à Regressão Simbólica de Funções	44

3.1.1	Considerações Iniciais	44
3.1.2	Representação do Indivíduo	46
3.1.3	Geração de uma População Inicial	47
3.1.4	Operador de Recombinação	48
3.1.5	Operadores de Mutação	48
3.1.6	Seleção de Indivíduos	49
3.1.7	Avaliação	49
3.1.8	Substituição	50
3.2	Descrição do Modelo de Solução de Problemas de Autovalores	51
3.2.1	Modelo Proposto assumindo Operadores Auto-Adjuntos Não-Negativos	51
3.2.1.1	Considerações Iniciais	51
3.2.1.2	Modelo de PG	52
3.2.1.3	Cálculo dos Demais Autovalores	56
3.2.2	Modelo Proposto assumindo Operadores Auto-Adjuntos Não-Positivos	58
3.2.3	Modelo Proposto assumindo Operadores Auto-Adjuntos Indefinidos	58
3.2.4	Modelo Proposto assumindo Operadores Normais	60
3.2.5	Propostas de Adaptações para Outros Operadores	62
3.2.6	Modelo de Solução Paramétrico	67
4	Estudos de Caso: Operadores Normais	69
4.1	Estudo de Caso: Operador de Laplace	69
4.1.1	Definição	69
4.1.2	Experimento: domínio retangular com solução analítica	70
4.1.3	Experimento: Domínio senoidal	77
4.1.3.1	Modelagem Proposta	79
4.1.3.2	Formulação por Método das Diferenças Finitas	81
4.1.3.3	Formulação por Método dos Elementos Finitos	83
4.1.3.4	Resultados Comparativos	84
4.2	Oscilador Harmônico Quântico (OHQ)	89
4.3	Conclusões do Capítulo	94

5	Estudos de Caso: Operadores Não Auto-Adjuntos	96
5.1	Descrição Preliminar do Problema	96
5.2	Formulação Matemática	97
5.2.1	Premissas deste modelo	97
5.2.2	Dedução do modelo eletromagnético	100
5.2.3	Dedução do Problema de Autovalores	104
5.3	Modelo semi-analítico de solução	107
5.4	Modelagem eletromagnética dos materiais	111
5.5	Resultados	112
5.5.1	Experimento não-paramétrico	114
5.5.2	Experimento paramétrico	119
5.6	Conclusões do Capítulo	124
6	Conclusões e Trabalhos Futuros	126
A	Produtos internos	135

Lista de figuras

Figura 1 - Grid de pontos utilizado pelo MDF.	33
Figura 2 - Domínio quadrado discretizado utilizando uma malha 6x6 de elementos quadrados.	37
Figura 3 - Representação com uma árvore para a função $f(x) = 2x + \text{sen}(x)$.	47
Figura 4 - Exemplo de cromossomo com $k_a = 3$ e $k_s = 1$. A função escalar f_1 é combinação linear das árvores A_1, A_2, A_3 , enquanto que f_2 é combinação linear de A_3, A_4, A_5 e f_3 é combinação linear de A_5, A_6 e A_7 .	53
Figura 5 - Média do melhor indivíduo dentre os 40 experimentos para cada quantidade de árvores para encontrar o menor autovalor.	73
Figura 6 - Domínio S das funções do operador de Laplace para o estudo de caso da seção 4.1.3. A senóide tem frequência $k = 6$.	78
Figura 7 - Domínio S com $k = 6$, $a = 1$ e $\alpha = 0.1$, discretizado para o Método Das Diferenças finitas, utilizando 15 pontos por direção (total: 15^2 pontos).	82
Figura 8 - Comparativo de acurácia entre os métodos para valores de k variando entre 15 e 105, com passo de 10.	86
Figura 9 - Comparação entre o tempo de execução dos métodos para atingir um erro menor ou igual ao um valor pré-estabelecido. O experimento foi executado com $k = 200$.	87
Figura 10 - Corte de um poço.	98
Figura 11 - Modelo de coluna de produção com escoamento em regime anular.	100
Figura 12 - Representação de um guia de ondas dividido em 4 camadas, sendo a última semi-infinita.	103

Figura 13- Comparação do tempo de 1 execução do método proposto baseado em PG com Q^2 execuções dos métodos clássicos para determinar o primeiro modo de propagação ($m = 1$) do poço de petróleo na frequência $f = 0,8$ GHz e óleo do tipo “D”. 120

Figura 14- Comparação do tempo de 1 execução do método proposto baseado em PG com Q^2 execuções dos métodos clássicos para determinar o primeiro modo de propagação ($m = 1$) do poço de petróleo na frequência $f = 1$ kHz e óleo do tipo “D”. 122

Lista de tabelas

Tabela 1 - Modelo proposto para achar os autopares do operador de Laplace.	72
Tabela 2 - Taxa de sucesso do modelo proposto para achar o menor autopar do operador de Laplace.	73
Tabela 3 - Taxa de sucesso do modelo proposto para achar o 2o. autopar do operador de Laplace.	74
Tabela 4 - Quantidade de rodadas cuja pontuação da solução é superior a determinado percentual para encontrar $f_{1,1}$ (colunas).	77
Tabela 5 - Parâmetros do modelo proposto para achar os autopares do operador de Laplace no domínio S .	85
Tabela 6 - Parâmetros utilizados pelo Método dos Elementos Finitos para achar os autopares do operador de Laplace no domínio S .	85
Tabela 7 - Parâmetros utilizados pelo Método das Diferenças Finitas para achar os autopares do operador de Laplace no domínio S .	85
Tabela 8 - Acurácia obtida resolvendo o problema paramétrico para achar o primeiro autopar do operador de Laplace no domínio S , utilizando $k = 200$.	88
Tabela 9 - Acurácia obtida resolvendo o problema paramétrico para achar o segundo autopar do operador de Laplace no domínio S , utilizando $k = 200$.	88
Tabela 10 - Modelo proposto para achar os autopares do oscilador harmônico quântico.	90
Tabela 11 - Resultados para a primeira autofunção do oscilador harmônico quântico.	91
Tabela 12 - Resultados para a primeira autofunção do oscilador harmônico quântico – 4000 gerações.	91

Tabela 13 - Resultados para a primeira autofunção do OHQ – número de rodadas que convergiram com erro $< 10^{-4}$ (células) para cada faixa de gerações (colunas).	92
Tabela 14 - Resultados para a 3a. menor autofunção do OHQ – número de rodadas que convergiram com erro $< 10^{-4}$ (células) para cada faixa de gerações (colunas).	92
Tabela 15 - Modelo proposto para achar o terceiro autopar do oscilador harmônico quântico (sem utilizar funções de Hermite).	93
Tabela 16 - Resultados para a terceira autofunção do OHQ – número de rodadas que convergiram com erro $< 10^{-4}$ (células) para cada faixa de gerações (colunas), experimento sem funções de Hermite.	93
Tabela 17 - Premissas da modelagem da coluna de produção	99
Tabela 18 - Constantes do meio de propagação “óleo”	112
Tabela 19 - Constantes dos meios de propagação	112
Tabela 20 - Modelo proposto para achar os modos do poço de petróleo.	114
Tabela 21 - Valor de κ_z para os primeiros modos com $\kappa_z \in [0, 1 - 150i; 100 - 10^{-4}i]$, frequência 0,8 GHz, óleo tipo “D”, por ordem crescente de atenuação.	114
Tabela 22 - Valor de κ_z para os primeiros modos com $\kappa_z \in [0, 1 - 150i; 100 - 10^{-4}i]$, frequência 0,8 GHz, óleo tipo “F”, por ordem crescente de atenuação.	115
Tabela 23 - Valor de κ_z para os primeiros modos com $\kappa_z \in [0, 1 - 150i; 100 - 10^{-4}i]$, frequência 0,8 GHz, óleo tipo “G”, por ordem crescente de atenuação.	115
Tabela 24 - Erro de cada um dos métodos para obtenção dos modos apresentados na tabela 21.	116
Tabela 25 - Erro de cada um dos métodos para obtenção dos modos apresentados na tabela 22.	117

Tabela 26 - Erro de cada um dos métodos para obtenção dos modos apresentados na tabela 23.	117
Tabela 27 - Valor de κ_z para os primeiros modos com $\kappa_z \in [10^{-7} - 150i; 100 - 10^{-7}i]$, frequência 1 kHz, óleo tipo “D”, por ordem crescente de atenuação.	118
Tabela 28 - Valor de κ_z para os primeiros modos com $\kappa_z \in [10^{-7} - 150i; 100 - 10^{-7}i]$, frequência 1 kHz, óleo tipo “F”, por ordem crescente de atenuação.	118
Tabela 29 - Erro de cada um dos métodos para obtenção dos modos apresentados na tabela 27.	118
Tabela 30 - Erro de cada um dos métodos para obtenção dos modos apresentados na tabela 28.	118
Tabela 31 - Erros médio, máximo, mínimo e desvio padrão do erro de cada um dos métodos para obtenção do primeiro modo – problema paramétrico para frequência de 0,8 GHz e óleo do tipo “D”.	121
Tabela 32 - Erros médio, máximo, mínimo e desvio padrão do erro de cada um dos métodos para obtenção do primeiro modo – problema paramétrico para frequência de 1 kHz e óleo do tipo “D”.	121
Tabela 33 - Valores numéricos da parte real dos coeficientes da equação 5-47. Coeficientes menores do que 10^{-7} foram considerados nulos para simplificar a exibição na tabela. O experimento usou óleo tipo “D” e frequência de 1 kHz.	123
Tabela 34 - Expressão analítica dos coeficientes κ_ρ^j da equação 5-47. O experimento usou óleo tipo “D” e frequência de 1 kHz.	123

*Mais cedo ou mais tarde, a teoria sempre
acaba assassinada pela experiência.*

Albert Einstein, *fonte desconhecida.*

1 Introdução

Diversos problemas em física e matemática modelados por equações diferenciais parciais resumem-se a encontrar o conjunto de autovalores e autofunções associadas. Alguns exemplos incluem a *busca de modos de propagação* em guias de ondas eletromagnéticas, a busca de *estados de excitação* em sistemas quânticos e a determinação de *frequências de vibração ressonantes* em estruturas mecânicas. Em particular, a motivação principal para o presente trabalho é a *pesquisa em telemetria sem fio de poços de petróleo*. Nesta, deseja-se estudar como o poço de petróleo comporta-se para propagar ondas eletromagnéticas, considerando os modos de escoamento do óleo (bolha ou anular), a presença de bifurcações e outras variações de geometria, poços com dupla completação, não-homogeneidades no óleo, incrustações de carbonato de cálcio (CaCO_3) e influência das formações rochosas.

Dependendo da aplicação, várias técnicas podem ser empregadas, dentre elas, aquelas baseadas em elementos finitos, diferenças finitas e solução semi-analítica das equações [1, 2, 3]. Tais métodos, em geral, empregam a discretização no domínio do operador diferencial para converter o problema original em um problema de encontrar autovalores de uma matriz finita, geralmente com muitos elementos. Tais métodos são eficazes, mas apresentam limitações associadas à representação de um operador de dimensão infinita por uma matriz finita, o que causa alto consumo de memória, perda de acurácia e a presença de falsas soluções próximas aos autovalores exatos. Além disso, se o domínio do operador for alterado de alguma forma, a simulação precisa ser executada novamente, porque a representação matricial do operador depende dos parâmetros do problema. Logo, aplicações que precisem avaliar a solução para um conjunto de parâmetros tipicamente demandam centenas ou milhares de simulações, o que torna o processo lento. Exemplos de aplicações deste tipo incluem otimização de parâmetros de sistemas físicos e síntese de sistemas robustos sob incertezas, como ocorre na análise de propagação em poços, que tem incertezas quanto à composição do óleo e quanto à sua temperatura.

Este trabalho investiga uma nova abordagem na qual técnicas de computação evolucionária são aplicadas para encontrar, simbolicamente, autovalores e autofunções de operadores diferenciais, geralmente vetoriais. Neste trabalho,

utiliza-se o termo operador vetorial para indicar uma transformação linear que atua sobre funções cujo contradomínio é o conjunto dos números complexos n -dimensional. Para isso, o modelo proposto aplica programação genética para evoluir as autofunções de maneira simbólica, reduzindo a um mínimo a carga de computação numérica. Como resultado, esquemas de discretização com alto nível de detalhe deixam de ser necessários e soluções espúrias passam a ser bem menos frequentes. Isto ocorre porque a computação simbólica, inerente ao modelo, reduz significativamente erros de arredondamento e de truncamento durante o processo. Além disso, como as autofunções são evoluídas em sua forma analítica, passa a ser possível incorporar ao processo de evolução da função parâmetros que representem características do domínio. Tais parâmetros tipicamente representam diâmetros, comprimentos ou outros parâmetros físicos característicos da aplicação. Conseqüentemente, a solução obtida pelo modelo proposto é válida para uma classe de problemas, diferentemente das soluções dadas pelas técnicas clássicas mencionadas, que demandam que o domínio seja completamente especificado. De maneira prática, tais soluções paramétricas são úteis em diversos cenários, tais como análises de sensibilidade, otimização de parâmetros de sistemas e estudo de incertezas.

O presente texto apresenta o modelo completo para operadores diferenciais normais, utilizando o conceito de indivíduos multi-árvore e empregando algoritmos de diferenciação e simplificação simbólica. O modelo evolui as autofunções e, a partir delas, calcula os autovalores empregando a *razão de Rayleigh*. O problema de encontrar as autofunções é definido como um problema de minimização, que é resolvido utilizando programação genética.

Este trabalho também mostra os resultados da aplicação do modelo em problemas clássicos com diversos níveis de complexidade. Os casos apresentados são a busca de autofunções do operador de Laplace, o oscilador harmônico quântico e a determinação de modos em guias de onda estratificados. Além disso, a presente proposta também mostra uma adaptação do modelo para casos particulares de operadores que não são normais. Como estudo de caso, deseja-se entender, utilizando o modelo proposto, como ocorre a propagação de ondas eletromagnéticas em um poço de petróleo, considerando diversas complexidades como não homogeneidades no meio, modos de escoamento do óleo e incertezas na composição dos fluidos.

Nos próximos capítulos, este texto foca em responder às seguintes perguntas: (1) qual é o problema abordado e qual é a sua relevância? (2) qual é a contribuição original deste trabalho? (3) qual é o estado atual da literatura relacionada? (4) qual é o modelo proposto para resolver o problema em questão? (5) quais são os resultados obtidos? (6) quais são as conclusões desta

pesquisa? Com o foco de responder claramente à estas perguntas, esta tese de Doutorado está estruturada como se segue.

O capítulo 2 destina-se a responder às questões (1), (2) e (3). Para isso, o capítulo apresenta uma formulação do problema de determinação de autopares de operadores diferenciais com exemplos. Além disso, o capítulo apresenta uma revisão literária, mostrando as técnicas disponíveis. Ademais, o capítulo também mostra como a programação genética têm sido aplicada em problemas semelhantes, como na solução de equações diferenciais parciais completamente especificadas por suas condições de contorno.

O capítulo 3 dedica-se à responder à pergunta (4). Para isso, ele descreve a técnica de programação genética e apresenta o modelo de solução de problemas de autopares de operadores diferenciais. O capítulo aborda primeiramente os operadores auto-adjuntos, pois estes apresentam propriedades interessantes que são exploradas pelo modelo, como a existência apenas de autovalores reais. Posteriormente, é apresentado o modelo de solução de problemas de autovalores para operadores normais, baseando-se no modelo para operadores auto-adjuntos. Então, o capítulo mostra outros tipos de problemas de autovalores, em particular o problema de autovalores generalizado e o problema de autovalores quadrático. Finalmente, o capítulo ainda mostra como o modelo pode ser adaptado para outros tipos particulares de operadores que não são normais, já preparando o terreno para a aplicação em análise de propagação em poços de óleo.

Na sequência, os capítulos 4 e 5 mostram os resultados obtidos, respondendo, portanto, à questão (5).

O capítulo 4 dedica-se a mostrar estudos de casos de problemas de determinação de autopares de operadores diferenciais vetoriais normais. O objetivo é mostrar como a técnica se adapta a cada problema e comparar os resultados com aqueles obtidos por técnicas clássicas. Além disso, o capítulo mostra também como os referidos problemas podem ser parametrizados, de modo que o modelo proposto entregue uma solução analítica e paramétrica.

O capítulo 5 dedica-se a ilustrar como o modelo proposto pode ser útil para prever a propagação de ondas eletromagnéticas em poços de petróleo. Para tal, o capítulo explica brevemente como o poço pode ser visto como um guia de ondas estratificado e formula um problema de determinação de autopares baseado em um operador que não é normal. Posteriormente, o capítulo apresenta uma solução semi-analítica e contrasta seus resultados com aqueles obtidos pelo modelo proposto neste trabalho. Finalmente, são feitas várias análises para diversas configurações de poço, variando tipo de óleo, tipo de escoamento e tipo de formação rochosa. Os resultados são dados em termos

das expressões analíticas dos campos elétricos, parametrizadas pelo diâmetro do poço e pelo percentual de incrustação de CaCO_3 .

Finalmente, o capítulo 6 apresenta as conclusões finais e propõe os próximos passos desta pesquisa. Logo, a pergunta (6) será respondida neste capítulo.

2

Descrição do Problema e Revisão da Literatura

O propósito deste capítulo é responder às seguintes perguntas:

1. qual é o problema abordado pela tese?
2. Qual é o estado atual da literatura?
3. Quais são as contribuições originais do presente trabalho de doutorado?

Para responder à primeira questão, este capítulo começa apresentando a formulação matemática geral para o problema abordado neste trabalho. Nas próximas seções, o capítulo define os operadores diferenciais vetoriais e mostra o conceito de autovalores e autofunções, também conhecidos pelo termo autopares.

Na sequência, o capítulo foca em responder à segunda questão. Para isso, o capítulo apresenta uma revisão da literatura sobre o assunto, mostrando as técnicas que mais têm sido usadas. O capítulo apresenta as técnicas clássicas baseadas em discretização de domínio, mais especificamente elementos finitos e diferenças finitas. Ademais, o capítulo também apresenta como a programação genética (PG) têm sido aplicada na literatura para resolver problemas semelhantes.

Finalmente, o presente capítulo conclui respondendo à terceira questão, isto é, resumindo quais são as contribuições originais.

2.1

Formulação Matemática de Problemas de Autovalores

O propósito desta seção é introduzir os elementos básicos de análise funcional, especialmente aqueles necessários para entender problemas de autovalores de operadores diferenciais. Ao final desta discussão, o leitor terá entendido os pré-requisitos básicos e entenderá formalmente qual é o problema abordado nesta tese.

Para isso, a ideia é apresentar sucintamente conceitos básicos de estruturas matemáticas, começando pela definição de conjuntos e evoluindo em grau de sofisticação até o conceito de operadores diferenciais com domínio em um *espaço de Hilbert*. Sempre que possível, trabalhar com *espaço de Hilbert* é interessante devido ao grande número de abstrações que os acompanha, como

os conceitos de conjuntos, métricas, vetores, soma vetorial e multiplicação por escalar, dimensão, norma, produto interno e ortogonalidade. Como consequência, os *espaços de Hilbert* exibem propriedades poderosas que tipicamente são exploradas em aplicações de *Física e Matemática*, como o conceito de completude. Este conceito, por exemplo, permite que seja possível representar qualquer elemento de um *espaço de Hilbert* a partir de uma combinação de outros elementos do mesmo espaço, como é o caso da *série de Fourier*. Todo conteúdo desta seção foi resumido a partir de [4, 5, 6].

2.1.1 Pré-requisitos Matemáticos

2.1.1.1 Conjuntos

Conjuntos são coleções de objetos distintos. Obviamente, faz parte desta definição um conceito abstrato de igualdade entre elementos desta coleção. Este texto utiliza os termos “conjunto” e “espaço” como sinônimos.

Dentre os conjuntos de interesse nesse trabalho, estão o conjunto dos números inteiros, \mathbb{Z} , o conjunto dos números reais, \mathbb{R} , e o conjunto dos números complexos, \mathbb{C} . Intervalos de números reais são representados por “[]”, ou seja, $[a; b]$ é o intervalo de todos os números reais entre a e b , com os extremos incluídos. No caso de a ou b serem números complexos, a notação $[a; b]$ representa o retângulo no plano complexo formado por todos os complexos z tais que $\text{Re } z \in [\text{Re } a; \text{Re } b]$ e $\text{Im } z \in [\text{Im } a; \text{Im } b]$, onde $\text{Re}(\cdot)$ e $\text{Im}(\cdot)$ representam, respectivamente, as partes real e imaginária.

Nem todos os conjuntos que serão abordados aqui são conjuntos numéricos. Em particular, este trabalho utiliza também conjuntos de funções, como o conjunto de todas as funções quadrado integráveis. Tais conjuntos serão introduzidos no texto quando for mais oportuno.

Finalmente, algumas operações sobre conjuntos são de interesse deste trabalho. São elas:

1. *interseção*: $A \cap B$ é o conjunto formado pelos elementos que pertencem simultaneamente à A e B ;
2. *união*: $A \cup B$ é o conjunto formado pelos elementos que pertencem a A ou B ;
3. *diferença*: $A - B$ é o conjunto formado pelos elementos que pertencem a A mas não pertencem à B ;

4. *produto cartesiano*: $A \times B$ é o conjunto formado por todos os pares ordenados (a, b) , $\forall a \in A$ e $\forall b \in B$.

2.1.1.2

Espaço Métrico

Espaço métrico é um espaço (ou conjunto) equipado de uma métrica. Intuitivamente, uma métrica é uma função que transforma dois elementos quaisquer do espaço em um número real positivo, de modo a medir o grau de diferença entre eles.

Formalmente, uma métrica de um espaço E é uma função $d(x, y)$, definida para todo par $(x, y) \in E \times E$ que satisfaz os axiomas:

1. $d(x, y) \geq 0$.
2. $d(x, y) = 0$ se, e somente se, $x = y$.
3. $d(x, y) = d(y, x)$.
4. $d(x, y) \leq d(x, z) + d(z, y)$, para todo $z \in E$ (desigualdade triangular).

Alguns exemplos de métrica incluem a função $d(x, y) = |x - y|$, definida no conjunto dos números complexos e a função $f(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})}$ para vetores do espaço N -dimensional \mathbb{R}^N , onde “ \cdot ” é o produto escalar (ver apêndice A).

2.1.1.3

Espaço Linear

Espaço linear (ou vetorial) é um espaço equipado de operações vetoriais. Formalmente, é necessário especificar um conjunto especial, denominado corpo, para poder definir um espaço linear. Sem entrar em detalhes técnicos, um corpo é um conjunto que define soma e produto entre seus elementos e que garante que os resultados destas operações pertençam ao corpo. Exemplos incluem o corpo dos números complexos, \mathbb{C} , o corpo dos números reais, \mathbb{R} , e o corpo dos números inteiros, \mathbb{Z} .

Dado o corpo C , um espaço linear L é definido como um conjunto de elementos que satisfaz uma série de propriedades, dentre elas:

1. Presença de uma operação, a soma vetorial “+”, que é capaz de transformar dois elementos quaisquer de L em outro elemento de L de maneira comutativa.
2. Presença de um elemento nulo, 0 , que satisfaz $v + 0 = v$ para todo $v \in L$.

3. Presença de um elemento $y \in L$ para todo $x \in L$, tal que $x + y = 0$. Neste caso, utiliza-se a notação $-x = y$. Além disso, usa-se a notação $x - z$ para abreviar $x + (-z)$, $z \in L$.
4. Se $\alpha \in C$, então existe uma operação de produto, conhecida como produto por escalar, que transforma todo elemento $v \in L$ em um elemento αv , que também pertence à L .
5. Existe um escalar, denotado por 1, tal que $1v = v$.

Existem outros axiomas usados na definição de espaço linear, mas uma definição formal completa foge ao escopo deste trabalho. É importante acrescentar, contudo, que o elemento α , $\alpha \in C$, é conhecido como escalar e o elemento v , $v \in L$, é conhecido como vetor.

Exemplos de espaços vetoriais conhecidos são o \mathbb{R}^N , definido usando o corpo dos números reais, e o \mathbb{C}^N , definido utilizando o corpo dos números complexos. Neste trabalho, a menos que seja especificado o contrário, estaremos trabalhando sempre com o corpo dos números complexos e, portanto, todos os escalares são complexos.

Espaços lineares são úteis para definir o conceito de independência linear. Formalmente, os k vetores $v_1, v_2, v_3, \dots, v_k$ são linearmente independentes se o único conjunto de escalares $(\alpha_1, \alpha_2, \dots, \alpha_k)$ capaz de fazer $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \dots + \alpha_k v_k = 0$ tem $\alpha_1 = \alpha_2 = \alpha_3 = \dots = \alpha_k = 0$. Se uma coleção de vetores não for linearmente independente, então diz-se que ela é linearmente dependente.

A partir da definição de independência linear, define-se a dimensão de um espaço vetorial. Basicamente, um espaço linear L é dito N -dimensional, ou tem N dimensões, se existem N elementos de L que são linearmente independentes, enquanto que $N + 1$ elementos são sempre linearmente dependentes. O espaço linear \mathbb{R}^N , por exemplo, é N dimensional, pois é possível selecionar no máximo N vetores linearmente independentes do \mathbb{R}^N .

Em alguns espaços, sempre será possível acrescentar um novo elemento a um conjunto finito de vetores linearmente independentes. Consequentemente, não existe número máximo de vetores linearmente independentes e, por isso, diz-se que tal espaço tem *dimensão infinita*. Um exemplo é o conjunto de todas as funções quadrado integráveis de suporte compacto.

Por fim, uma base para um espaço linear L de dimensão n é qualquer conjunto de n vetores linearmente independentes de L . As bases tem uma propriedade muito importante: qualquer vetor de L pode ser escrito como uma combinação linear dos vetores da base. Existem conceitos equivalentes para espaços de dimensão infinita, que serão abordados mais à frente neste texto.

2.1.1.4

Normas e Produtos Internos

Normas são funções semelhantes às métricas, mas elas necessitam de um espaço linear para serem definidas e, além disso, têm apenas 1 argumento, em oposição aos 2 argumentos da função métrica. Intuitivamente, as normas são uma medida de “distância ao vetor nulo”.

Formalmente, a norma de um vetor x do espaço linear L , denotada por $\|x\|$, é um número real não-negativo que satisfaz:

1. $\|x\| \geq 0$ e $\|x\| = 0$ se, e somente se, $x = 0$;
2. $\|\alpha x\| = |\alpha| \|x\|$, para todo escalar α ;
3. $\|x + y\| \leq \|x\| + \|y\|$, para todos os vetores x e y do espaço linear L .

Uma propriedade interessante das normas é que a função $d(x, y) = \|x - y\|$ satisfaz todas as propriedades de métrica e, por isso, é conhecida como métrica induzida pela norma. Um espaço linear equipado de uma norma é chamado de espaço linear normado.

Os produtos internos, por outro lado, são funções que operam sobre dois elementos do espaço linear, gerando um número complexo, em oposição à norma, que é sempre um número real não-negativo. Formalmente, o produto interno de dois vetores x e y de um espaço linear L , denotado por $\langle x; y \rangle$, é uma função complexa que satisfaz:

1. $\langle x; x \rangle \geq 0$ e $\langle x; x \rangle = 0$ se, e somente se, $x = 0$;
2. $\langle x; y \rangle = \overline{\langle y; x \rangle}$, onde a barra superior indica o complexo conjugado;
3. $\langle \alpha x; y \rangle = \alpha \langle x; y \rangle$, onde α é um escalar;
4. $\langle x + y; z \rangle = \langle x; z \rangle + \langle y; z \rangle$, para todo $z \in L$.

Espaços lineares equipados de produto interno são chamados de espaços produto interno. De maneira interessante, a função $f(x) = \sqrt{\langle x; x \rangle}$ apresenta todas as propriedades de norma e, por isso, é chamada de norma induzida pelo produto interno. Como resultado, todo espaço linear com produto interno também é um espaço linear normado.

Exemplos de produto interno incluem o produto escalar do \mathbb{R}^N . O apêndice A exemplifica uma série de produtos internos utilizados neste trabalho.

2.1.1.5

Espaço de Hilbert

Espaço de Hilbert é um conceito complexo de análise funcional, que envolve a ideia de completude. Explicar o conceito de completude neste texto foge ao escopo, mas intuitivamente diz-se que um espaço é completo quando todas as sequências convergentes formadas por elementos do espaço convergem para um elemento também do espaço. Note, contudo, que este conceito conforme apresentado é informal.

Dado o conceito de completude, o espaço de Hilbert é definido como um espaço linear com produto interno e completo. O espaço de Hilbert de maior interesse neste trabalho é o espaço formado por todas as funções de “energia finita” cujo domínio é o conjunto Ω . A notação usada para este espaço é $L^2(\Omega)$. Frequentemente, o contradomínio destas funções é um vetor do espaço vetorial \mathbb{C}^N , e, neste caso, usa-se a notação $L^2(\Omega)^N$. Tipicamente, $N = 2$ ou $N = 3$.

Sendo assim, se $f \in L^2(\Omega)^N$, então a função f satisfaz às seguintes propriedades:

1. o domínio de f é o conjunto Ω . Tipicamente, Ω é um subconjunto do \mathbb{R}^N ;
2. o contradomínio de f é o \mathbb{C}^N ;
3. f tem energia finita.

Para definir o conceito de energia finita, usa-se o produto interno. Basicamente, f tem energia finita se $\langle f; f \rangle \leq \infty$, ou seja, f tem norma finita. Neste caso, o produto interno é definido como:

$$\langle f_1; f_2 \rangle = \int_{\Omega} f_1(v) \cdot \overline{f_2(v)} w(v) dv, \quad (2-1)$$

onde f_1 e f_2 são funções de domínio Ω e contradomínio \mathbb{C}^N , w é uma função real positiva, conhecida como peso, e “ \cdot ” é o produto escalar conforme definido para vetores reais (ver apêndice A). Note que a equação 2-1 tem uma integral tipicamente multivariada, uma vez que, usualmente, Ω é um espaço linear de dimensão maior do que 1.

Outros espaços de Hilbert de interesse são o \mathbb{R}^N equipado com o produto interno tradicional $\langle \mathbf{x}; \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y}$ e o \mathbb{C}^N , com o produto interno $\langle \mathbf{x}; \mathbf{y} \rangle = \mathbf{x} \cdot \overline{\mathbf{y}}$, onde a barra indica o complexo conjugado.

2.1.1.6**Ortogonalidade**

O conceito de ortogonalidade está diretamente relacionado ao conceito de produto interno. Dois vetores x e y do espaço de Hilbert H são ortogonais se $\langle x; y \rangle = 0$. Se $\|x\| = \sqrt{\langle x; x \rangle}$, então, para quaisquer vetores ortogonais x e y , $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

Um conjunto de vetores de um espaço de Hilbert é dito *ortogonal* quando todos os vetores deste conjunto são ortogonais dois a dois. Se, além disso, os vetores tiverem norma igual a 1, então o conjunto é dito *ortonormal*.

Se um conjunto enumerável e ortonormal de vetores $\{x_n\}_{n \geq 1}$ de um espaço de Hilbert H for tal que qualquer vetor v de H pode ser escrito como

$$v = \sum_n \langle v; x_n \rangle x_n, \quad (2-2)$$

então o conjunto $\{x_n\}_{n \geq 1}$ é uma *base (de Shaulder)* ortonormal para H .

Alternativamente, se $\{x_n\}_{n \geq 1}$ for uma base ortonormal para o espaço de Hilbert H , então para todo $\epsilon > 0$, existe um inteiro positivo N_ϵ tal que, para todo $N \geq N_\epsilon$,

$$\left\| v - \sum_{n=1}^N \langle v; x_n \rangle x_n \right\| \leq \epsilon. \quad (2-3)$$

Sendo assim, para valores suficientemente grandes de N , sempre é possível aproximar um vetor v com erro arbitrariamente pequeno utilizando os elementos da base de Shaulder ortonormal $\{x_n\}_{n \geq 1}$. Na prática, busca-se bases que demandem valores pequenos de N_ϵ , de modo a reduzir a carga computacional.

2.1.1.7**Operadores Lineares**

Neste trabalho, operadores lineares, também conhecidos como transformações lineares, são funções que operam em espaços de Hilbert e satisfazem o princípio da superposição. Formalmente, se $f : H_1 \rightarrow H_2$ é uma função que transforma elementos do espaço de Hilbert H_1 em elementos do espaço de Hilbert H_2 , então f é linear se $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$, para quaisquer escalares α, β e quaisquer vetores x e y de H_1 . Note que a definição faz sentido apenas se o corpo utilizado para definir H_1 e H_2 for o mesmo.

Exemplos simples de operadores lineares incluem a função real $f(x) = ax$, $a \in \mathbb{R}$ e o operador conjugado, $f(z) = \bar{z}$, $z \in \mathbb{C}$, desde que definido num espaço de Hilbert cujo corpo é o conjunto dos números reais. Outros casos interessantes incluem o operador derivada e o operador integral.

Dada a definição de operador linear, este trabalho utiliza o termo *operador diferencial linear* para se referir a qualquer operador linear que

contenha derivadas. Geralmente, estes operadores são definidos para elementos de $L^2(\Omega)^N$, e exemplos típicos incluem a derivada tradicional, o laplaciano, o rotacional e o divergente. Este trabalho entrará em detalhes acerca destes operadores mais adiante, quando for necessário.

2.1.1.8

Operadores Limitados

Operadores limitados são aqueles cuja norma do valor transformado por eles não “explode”. De maneira formal, se A é um operador linear limitado de H_1 em H_2 , então existe um $k > 0$ tal que $\|Ax\|_2 \leq k \|x\|_1$ para todo $x \in H_1$. Nessa definição, $\|\cdot\|_1$ e $\|\cdot\|_2$ são as normas de H_1 e H_2 respectivamente. Se um operador não for limitado, ele é chamado de *operador ilimitado*.

Em geral, operadores definidos em termos de derivadas são ilimitados. Por outro lado, operadores definidos em termos de integrais são geralmente limitados.

2.1.1.9

Adjuntos de Operadores Lineares

O operador adjunto é um conceito semelhante ao conceito de matriz transposta conjugada ou hermitiana. Uma definição formal geral de operador adjunto envolve o conceito de operadores densamente definidos, que foge ao escopo deste texto. Contudo, informalmente, se A é um operador linear entre os espaços de Hilbert H_1 e H_2 , então o adjunto de A , representado pela notação A^* , é um operador linear que satisfaz $\langle Ax; y \rangle_2 = \langle x; A^*y \rangle_1$ para todo $x \in H_1$ e $y \in H_2$, onde $\langle \cdot; \cdot \rangle_1$ e $\langle \cdot; \cdot \rangle_2$ são os produtos internos de H_1 e H_2 . Note que H_2 passa a ser o domínio do operador adjunto.

No caso de operadores limitados, essa definição informal é, na realidade, a definição precisa de operador adjunto. Para operadores ilimitados densamente definidos, esta também é a definição formal, mas não é totalmente rigorosa, porque o domínio do adjunto é, em geral, um subconjunto de H_2 .

Seja A um operador definido de H em H . Se o adjunto de A é ele mesmo, então A é chamado de *operador auto-adjunto*. Caso A comute com o seu adjunto, tal operador é chamado de *operador normal*. Sendo assim, se $AA^*x = A^*Ax$ para todo $x \in H$, então A é um operador linear normal. Operadores normais e operadores auto-adjuntos são de interesse deste trabalho por apresentarem correlação com problemas físicos de interesse. Ademais, tais classes de operadores possuem propriedades interessantes que serão exploradas no modelo de solução de problemas de autovalores que é cerne desta pesquisa.

2.1.1.10

Operadores Positivos e Negativos

O conceito de matrizes positivas e negativas é facilmente extensível à operadores em espaços de Hilbert. Seja A um operador linear, auto-adjunto, de H para H e seja $\langle \cdot; \cdot \rangle$ o produto interno. Temos as seguintes possibilidades:

1. se $\langle Ax; x \rangle \geq 0$ para todo $x \in H$, então A é um operador *não-negativo*;
2. se $\langle Ax; x \rangle > 0$ para todo $x \in H$, então A é um operador *positivo*;
3. se existir $k > 0$ tal que $\langle Ax; x \rangle > k \|x\|^2$ para todo $x \in H$, então A é um operador *positivo definido*.

Da mesma forma, é possível definir operadores não-positivos, negativos e negativos definidos, apenas invertendo as desigualdades. Note que tais definições são válidas apenas para operadores auto-adjuntos.

Quando um operador auto-adjunto não satisfaz a nenhuma dessas classificações, então ele é *indefinido*.

2.1.2

Problemas de Autovalores de Operadores Diferenciais Lineares

2.1.2.1

Considerações Preliminares

Esta seção dedica-se a explicar o principal problema abordado por esta pesquisa: a busca de autopares de operadores diferenciais lineares. Um autopar é um par de autovalor e autovetor, que são objetos com propriedades especiais de um operador linear. De maneira simplificada, quando um dado operador não modifica um vetor específico (a menos uma multiplicação por escalar), então este vetor é um autovetor do operador. Autovetores são quantidades importantes na física, pois estão geralmente relacionados à grandezas físicas de interesse, como modos de propagação em guias de onda eletromagnética, estados de excitação em sistemas quânticos e frequências de vibração em estruturas mecânicas.

Esta seção apresenta 3 classificações diferentes para problemas de autovalores. Seja A um operador linear definido num espaço de Hilbert H , isto é, H é o domínio e o contradomínio de A . Seja λ um escalar do corpo de H e seja x um vetor de H .

O **problema de autovalores clássico** consiste em encontrar as soluções (caso existam) da equação

$$Ax = \lambda x. \quad (2-4)$$

Se x' e λ' são soluções de 2-4, então x' é um autovetor de A , λ' é o autovalor associado ao autovetor x' e o par (λ', x') é um autopar de A . O conjunto de autopares de um operador linear pode ser vazio (nenhum autovalor), finito ou infinito. Neste último, o conjunto ainda pode ser enumerável (discreto) ou não-enumerável.

Se B é um operador linear também definido em H , então o **problema de autovalores generalizado** consiste em encontrar as soluções da equação

$$Ax = \lambda Bx. \quad (2-5)$$

Neste caso, se x' e λ' são soluções de 2-5, então x' é um autovetor generalizado de A com relação a B , λ' é o autovalor associado ao autovetor x' e o par (λ', x') é um autopar de A com relação a B .

Muitas vezes, um problema prático não pode ser escrito nas duas formas citadas acima. Por isso, é conveniente generalizar os problemas de autovalores, de modo que a forma mais geral englobe todos (ou a maioria) dos casos práticos. O **problema de autovalores não-padrão** (*nonstandard eigenvalue problem*) consiste em encontrar as soluções da equação

$$A(\lambda) x = 0. \quad (2-6)$$

Neste caso, o operador linear A é função do autovalor λ , ou seja, para cada valor distinto de λ tem-se um operador linear $A(\lambda)$. Mais uma vez, se x' e λ' são soluções de 2-6, então (λ', x') é um autopar (não-padrão) de A .

Note que tanto o problema clássico de autovalores (equação 2-4) quanto o problema generalizado (equação 2-5) se reduzem ao problema não-padrão (equação 2-6). No primeiro, temos $U(\lambda) = (A - \lambda I)$, onde I é o operador identidade, e resolve-se na forma não padrão $U(\lambda) x = 0$. No segundo caso, temos $U(\lambda) = (A - \lambda B)$.

Esta tese foca no problema padrão, apresentado em 2-4. Sendo assim, deste parágrafo em diante este texto refere-se sempre ao problema de autovalores padrão, da equação 2-4. No entanto, também são abordados alguns casos particulares do problema de autovalores não padrão, como o problema de autovalores quadrático, no qual $U(\lambda) = A_0 + \lambda A_1 + \lambda^2 A_2$, onde A_0 , A_1 e A_2 representam operadores lineares definidos no espaço de Hilbert H .

Existem algumas propriedades interessantes para autopares de um operador linear [4, 6, 5]. Uma delas é a independência linear entre autovetores que correspondem a autovalores distintos. Mais especificamente, se (λ_1, x_1) , (λ_2, x_2) , \dots , (λ_n, x_n) são autopares de um operador A e os autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$ são distintos, então o conjunto $\{x_1, x_2, \dots, x_n\}$ é linearmente independente [5].

Outras propriedades de interesse dependem de conhecimento mais específico do operador. Por exemplo, operadores auto-adjuntos têm *autovalores reais*. Além disso, seus *autovetores correspondentes a autovalores distintos são ortogonais*. Ademais, se o operador for não-negativo, então todos os seus autovalores são não-negativos e se o operador for positivo, então todos os seus autovalores são positivos. Analogamente, se o operador for não-positivo, então todos os seus autovalores são não-positivos e se o operador for negativo, então todos os seus autovalores são negativos.

2.1.2.2

Razão de Rayleigh

Seja A um operador linear definido num espaço de Hilbert H e seja (λ, x') um autopar de A , ou seja, $Ax' = \lambda x'$. Assim, pode-se escrever

$$\lambda = \frac{\langle Ax'; x' \rangle}{\langle x'; x' \rangle}. \quad (2-7)$$

Portanto, dado o autovetor, o operador e o produto interno, é possível calcular o autovalor associado. A razão de Rayleigh é apenas uma generalização de 2-7 para o caso em que x' não necessariamente é um autovetor de A , ou seja, para cada $x \in H$, define-se

$$R(x) = \frac{\langle Ax; x \rangle}{\langle x; x \rangle}. \quad (2-8)$$

A função $R(x)$ definida em 2-8 é a *razão de Rayleigh*. Claramente, no caso particular em que x é um autovetor de A , 2-8 é equivalente a 2-7.

Para diversos tipos de operadores, a razão de Rayleigh é estacionária para valores de x iguais a autovetores de A . Informalmente, isso quer dizer que erros de primeira ordem na aproximação de um autovetor x' implicam em erros de segunda ordem (menores) para aproximar seu autovalor. Em particular, esta propriedade é verdadeira para operadores normais.

Se A for auto-adjunto, não-negativo e seu conjunto de autovalores for discreto, então o menor autovalor de A pode ser obtido encontrando o mínimo de $R(x)$. Analogamente, se A for não-positivo, então o maior autovalor de A é igual ao máximo de $R(x)$. Essas propriedades são úteis para a metodologia que será apresentada neste texto para encontrar autopares de operadores auto-adjuntos.

Esta seção tratou da definição dos problemas de autovalores. Nas próximas seções, este texto apresenta uma revisão das técnicas existentes para resolver tais problemas.

2.2

Técnicas para Solução de Problemas de Autovalores na Literatura

2.2.1

Considerações Iniciais

Esta seção foca em revisar as técnicas presentes na literatura para solução de problemas de autovalores de operadores lineares.

As técnicas disponíveis tipicamente aplicam discretização no domínio das funções para converter o operador de dimensão infinita em uma matriz. Posteriormente, tais técnicas aplicam algoritmos numéricos para encontrar os autovalores desta matriz. Como a matriz é finita, apenas um conjunto finito de autovalores é encontrado. Caso mais autovalores sejam necessários, é preciso discretizar o domínio com mais pontos, de modo a gerar matrizes maiores e, conseqüentemente, com mais autovalores. Além disso, dependendo da técnica, as autofunções são encontradas de formas variadas, todas elas baseadas em algum tipo de expansão em bases de funções ortogonais.

Dentre as principais técnicas, destaca-se o Método das Diferenças Finitas, o Método dos Elementos Finitos [7, 8] e o Método dos Volumes Finitos [9]. Para problemas de autovalores de operadores lineares envolvendo integrais, destacam-se as técnicas que se baseiam em discretização do contorno, como o Método dos Momentos [10] e o Método dos Elementos de Contorno (*boundary element method*) [11]. Por fim, ainda existem técnicas que não dependem da existência de malhas ou discretização do domínio, como os Métodos Espectrais [12].

O Método das Diferenças Finitas e o Método dos Elementos Finitos são geralmente os mais adequados para modelar geometrias sofisticadas, devido à presença de uma malha que discretiza o domínio. Tal malha é projetada para adaptar-se ao formato do domínio. Além disso, as técnicas baseadas em discretização de contorno – Método dos Momentos e outras – operam de maneira semelhante, logo são adequadas para problemas que têm contornos complexos e são formulados em termos de integrais. Estas técnicas usualmente têm uma grande carga computacional, pois geram matrizes muito grandes e, portanto, necessitam de muita memória e processamento, o que as torna inadequadas para geometria pouco complexas.

Por outro lado, as técnicas baseadas em soluções analíticas ou semi-analíticas costumam adequar-se bem a problemas com geometrias pouco complexas, porque modelam de maneira precisa tais geometrias utilizando figuras geométricas conhecidas (retângulos, cubos, cilindros, esferas, cones, etc). Tais técnicas são computacionalmente eficientes, mas pouco flexíveis.

A presente seção foca apenas nos métodos de elementos finitos e diferenças finitas. Como não é foco desta tese abordar operadores baseados em integrais, optou-se por deixar de fora uma revisão acerca de técnicas baseadas

em discretização de contorno. Adicionalmente, esta seção também não aborda as técnicas simbólicas porque estas dependem demasiadamente do problema ao qual elas são aplicadas. Por isso, as técnicas simbólicas e semi-simbólicas serão abordadas mais a frente, nos estudos de caso específicos, como objeto de comparação.

2.2.2

Método das Diferenças Finitas

O *Método das Diferenças Finitas* (MDF) é um dos métodos baseados em aproximação numérica utilizados na literatura para resolver equações diferenciais parciais [13, 14]. O domínio contínuo das funções é aproximado por uma malha finita de pontos e uma função no domínio original é representada por sua versão discreta, caracterizada pelos valores da função original em cada um dos pontos da malha.

Nesta representação discreta, a derivada em um ponto da malha é aproximada utilizando uma combinação linear dos valores da função em pontos adjacentes da malha. Assim, a versão discreta de um operador diferencial é escrita em termos de um conjunto de combinações lineares envolvendo os valores da função discreta na malha, sendo uma combinação para cada ponto da malha [7, 8, 13, 14].

A figura 1 mostra uma malha retangular uniforme, na qual deseja-se estimar a derivada de uma função $f(x, y)$ no ponto da malha de coordenadas (x_j, y_j) . Para isso, o MDF utiliza os pontos adjacentes. Para estimar a derivada na direção x , usa-se os pontos adjacentes com o mesmo valor da coordenada y . Analogamente, para estimar a derivada em y , utiliza-se os pontos adjacentes com o mesmo valor de x . Adotando a notação simplificada $f_{i,j} = f(x_i, y_j)$, pode-se escrever, por exemplo,

$$f_{[x]}(x_i, y_j) = \frac{-f_{i-1,j} + f_{i+1,j}}{2h} + \epsilon \quad (2-9)$$

$$f_{[y]}(x_i, y_j) = \frac{-f_{i,j-1} + f_{i,j+1}}{2h} + \epsilon. \quad (2-10)$$

Nas equações, o subscrito “[t]” indica a derivada na variável t . Além disso, h é a distância entre dois pontos adjacentes do grid numa mesma direção ($h = x_{i+1} - x_i = y_{j+1} - y_j$) e ϵ é o erro de aproximação.

As aproximações de diferenças finitas derivam da aplicação da série de Taylor gerada em torno do ponto que se deseja estimar a derivada. Ao truncar a série e avalia-la em uma quantidade suficiente de pontos adjacentes, é possível montar um sistema linear de equações no qual as derivadas no ponto de interesse do grid são as incógnitas. A solução, então, é a expressão de

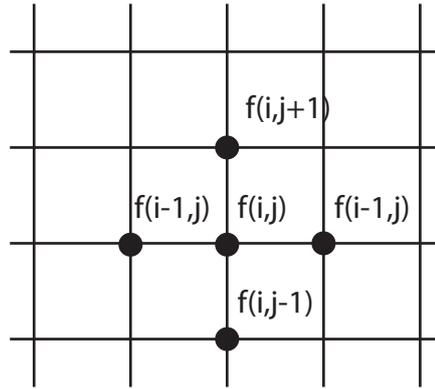


Figura 1: Grid de pontos utilizado pelo MDF.

aproximação da derivada. Quanto mais termos forem utilizados na série de Taylor, melhor é a aproximação da derivada, contudo mais pontos adjacentes serão usados. Na aproximação 2-10, por exemplo, o erro é proporcional a h^2 . Uma aproximação da primeira derivada com erro proporcional a h^4 seria:

$$f_{[y]}(x_i, y_j) = \frac{1}{h} \left(\frac{1}{12} f_{i,j-2} - \frac{2}{3} f_{i,j-1} + \frac{2}{3} f_{i,j+1} - \frac{1}{12} f_{i,j+2} \right) + \epsilon. \quad (2-11)$$

Note que, se $0 \leq h < 1$, então esta última aproximação gera menos erro pois, neste caso, $h^4 < h^2$. Em contrapartida, esta aproximação utiliza mais pontos da malha, o que tende a gerar sistemas de equações menos esparsos, que consomem mais memória e são mais difíceis de resolver.

Caso seja necessário encontrar a derivada na direção de um vetor unitário $\hat{\mathbf{n}} = (n_x; n_y)$, basta calcular o produto escalar deste e o gradiente de f , $\nabla f = (f_{[x]}; f_{[y]})$. Assim, a derivada de f na direção $\hat{\mathbf{n}}$ é igual a $f_{[\hat{\mathbf{n}}]} = (n_x f_{[x]}; n_y f_{[y]})$ e, usando 2-9 e 2-10, temos

$$f_{[\hat{\mathbf{n}}]} = n_x \frac{-f_{i-1,j} + f_{i+1,j}}{2h} + n_y \frac{-f_{i,j-1} + f_{i,j+1}}{2h} + \epsilon. \quad (2-12)$$

Para resolver um problema de autovalores utilizando o MDF, deve-se primeiramente utilizar as aproximações das derivadas expressas como diferenças, conforme mostrado nos últimos parágrafos. Posteriormente, aplica-se as condições de contorno do problema para, então, montar um sistema de equações lineares cuja solução é o valor de f em cada ponto do grid. Como um parâmetro desta equação é desconhecido, geralmente tem-se um problema de autovalores matricial. Os próximos parágrafos irão ilustrar um exemplo clássico.

Considere o operador de Laplace, definido para funções com domínio

bidimensional como

$$\begin{aligned}\nabla^2 f &= f_{[xx]} + f_{[yy]} \\ f &\in L^2([0; 1]^2) \\ f(0, y) &= f(1, y) = 0 \\ f(x, 0) &= f(x, 1) = 0 \\ x, y &\in [0; 1].\end{aligned}\tag{2-13}$$

Ou seja, o domínio do operador é o espaço de Hilbert formado por todas as funções complexas, definidas no retângulo $[0; 1]^2$, nulas no contorno e quadrado integráveis de acordo com o produto interno 2-1. Por conveniência, assume-se $w(v) = 1$ em 2-1.

A partir da definição do operador de Laplace, pode-se formular o problema de autovalores

$$\nabla^2 f = \lambda f,\tag{2-14}$$

que pode ser resolvido utilizando a aproximação da segunda derivada do MDF,

$$f_{[xx]}(x_i, y_j) = \frac{f_{i-1,j} - 2f_{i,j} + f_{i+1,j}}{h^2} + \epsilon\tag{2-15}$$

$$f_{[yy]}(x_i, y_j) = \frac{f_{i,j-1} - 2f_{i,j} + f_{i,j+1}}{h^2} + \epsilon.\tag{2-16}$$

Ao aplicar as aproximações 2-15 e 2-16 na definição 2-13, tem-se a versão discreta do operador de Laplace para o grid aqui exemplificado:

$$(\nabla^2 f)_{i,j} = \frac{1}{h^2} (f_{i+1,j} + f_{i-1,j} + f_{i,j+1} + f_{i,j-1} - 4f_{i,j}).\tag{2-17}$$

Sendo assim, ao substituir esta versão discreta do operador de Laplace em sua definição 2-13, obtêm-se a equação

$$f_{i+1,j} + f_{i-1,j} + f_{i,j+1} + f_{i,j-1} - 4f_{i,j} = h^2 \lambda f_{i,j},\tag{2-18}$$

válida para cada $(i; j)$ do grid, com $i = 2, 3, \dots, M - 1$ e $j = 2, 3, \dots, M - 1$, onde M é o número de pontos numa direção.

Como, pela definição do operador, f é nula no contorno, deve-se acrescentar essas equações ao sistema de equações 2-18. Se o grid tiver M pontos em cada direção, isso significa que

$$\begin{aligned}f_{1,j} &= f_{M,j} = 0, \quad j = 1, 2, \dots, M \\ f_{i,1} &= f_{i,M} = 0, \quad i = 1, 2, \dots, M.\end{aligned}\tag{2-19}$$

Considerando apenas as equações úteis para 2-18, temos um total de $(M - 2)^2$ variáveis e equações. As variáveis são os valores de $f_{i,j}$, $i, j = 1, 2, \dots, M - 1$

e as equações são dadas por 2-18, usando como complemento as condições de contorno 2-19.

Com isso, finalmente obtém-se um sistema de equações do tipo

$$C\mathbf{f} = h^2\lambda\mathbf{f}, \quad (2-20)$$

onde C é uma matriz de coeficientes numéricos e

$$\mathbf{f} = [f_{2,2} \quad f_{2,3} \quad \dots \quad f_{2,M-1} \quad f_{3,2} \quad f_{3,3} \quad \dots \quad f_{M-1,M-1}]^T, \quad (2-21)$$

é o vetor coluna formado pelas variáveis. A equação 2-20 é um problema de autovalores matricial, onde \mathbf{f} são os autovetores e $h^2\lambda$ são os autovalores.

Note que essa abordagem gera um número finito de autovalores distintos, no máximo igual à ordem da matriz C , i.e., $M - 2$. No entanto, o operador de Laplace tem infinitos autovalores, logo, caso mais autovalores sejam necessários, é preciso aumentar o número de pontos utilizados na discretização do domínio. Sendo assim, uma *importante limitação* desta técnica é que, para conseguir obter um número maior de autovalores, necessariamente é preciso trabalhar com matrizes maiores, o que aumenta o consumo de memória e o tempo de execução.

Outra limitação surge quando o domínio é infinito. Ainda no exemplo do operador de Laplace, considere o caso em que a função f está definida em todo o \mathbb{R}^2 , ou seja, ao invés da definição em 2-13, temos

$$\begin{aligned} \nabla^2 f &= f_{[xx]} + f_{[yy]} \\ f &\in L^2(\mathbb{R}^2) \\ f(t, y) &= 0 \text{ quando } t \rightarrow \pm\infty \\ f(x, v) &= 0 \text{ quando } v \rightarrow \pm\infty \\ x, y &\in \mathbb{R}. \end{aligned} \quad (2-22)$$

Neste caso, é impossível discretizar o domínio de modo a incluir o contorno, a menos que algum artifício seja adotado. Dependendo da aplicação, existem algumas formas de contornar essa questão.

Uma possibilidade é simplesmente trabalhar com uma malha com número de pontos suficientemente alto de modo que seja possível assumir que $f = 0$ nos limites da malha. Dependendo da aplicação, pode ser necessário adotar uma malha muito grande e, conseqüentemente, a matriz C teria muitos elementos, o que aumentaria demasiadamente o consumo de memória. Outra forma de contornar a questão é utilizando condições de contorno especiais nos limites da malha, que indiretamente garantam que a função f seja nula no infinito. Um exemplo é um conjunto de técnicas chamado *perfectly matched layer* (PML),

usado muito em eletromagnetismo [15, 16].

Um terceiro ponto negativo da técnica é que *as autofunções são dadas em sua forma discreta*, ou seja, seus valores são conhecidos apenas nos pontos da malha. Caso seja necessário avaliar a função fora dos pontos da malha, é preciso utilizar algum tipo de interpolação. Portanto, para aplicações que necessitem da função contínua original, o MDF não é o mais adequado. Neste caso, a técnica numérica mais adequada é o método dos elementos finitos (MEF), descrito a seguir.

2.2.3

Método dos Elementos Finitos

O *Método dos Elementos Finitos* (MEF) é considerado o mais geral e bem conhecido método de solução de equações diferenciais parciais (EDP). O MEF busca aproximar a função original por meio de funções com um certo grau de suavidade em todo o domínio, mas que são polinomiais por partes em células simples [17, 18]. Usualmente, tais células são triangulares ou retangulares, no caso bidimensional, e tetraedros no caso tridimensional. Contudo, existem aplicações com células mais complexas, inclusive com formatos curvos. Idealmente, o formato das células deve equilibrar a qualidade na representação do domínio com a simplicidade na montagem do sistema de equações final.

A ideia fundamental do MEF é aproximar a solução exata f utilizando aproximações válidas em sub-domínios, conhecidos como células ou elementos. A união destes sub-domínios é igual ao domínio completo e, portanto, a função f é aproximada em todos os pontos, diferentemente do MDF, que aproxima f num sub-conjunto finito de pontos do domínio. Geralmente, a função f é escrita por meio de uma interpolação polinomial dentro da célula e condições de contorno são aplicadas entre células vizinhas. Dentro de cada célula, o operador diferencial é aplicado sobre a representação aproximada de f , o que elimina as derivadas e, conseqüentemente, converte o operador diferencial em um operador algébrico.

Para realizar a interpolação no interior de cada célula, é preciso utilizar pontos de referência. Estes incluem os vértices das células e, caso necessário, outros pontos tomados das arestas das células. Estes pontos, então, passam a ser variáveis cujo o valor da função f é desconhecido, assim como no MDF. Sendo assim, o MEF também gera um sistema linear de equações cujas variáveis são os valores da função f nos pontos da malha. No entanto, por causa da existência de uma interpolação, o resultado do MEF é um conjunto de funções aproximadas que juntas podem ser usadas para estimar a solução exata f em qualquer ponto do domínio. Note, também, o MEF não requer

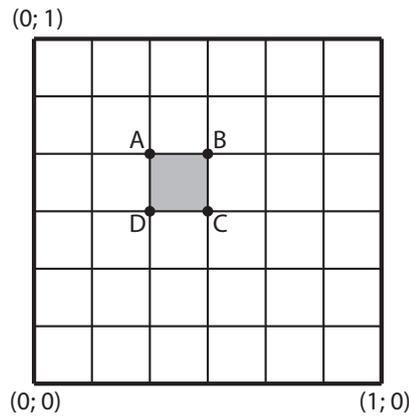


Figura 2: Domínio quadrado discretizado utilizando uma malha 6x6 de elementos quadrados.

aproximações especiais para as derivadas, uma vez que a expressão analítica de f é conhecida de maneira aproximada.

Voltemos ao exemplo de encontrar autovalores de 2-13. Uma possível solução por meio de elementos finitos poderia partir de uma discretização do domínio $[0; 1]^2$ em M células quadradas de lado $l = 1/M$ (ver figura 2). Portanto, a malha discreta seria dotada de M^2 quadrados. Queremos encontrar os valores da função exata f que satisfaz 2-13 nos vértices destes quadrados e, posteriormente, utilizar uma interpolação polinomial para calcular f em cada ponto do quadrado.

Considere um dado quadrado cujos vértices têm coordenadas $A = (x_1, y_1)$, $B = (x_2, y_2)$, $C = (x_2, y_1)$ e $D = (x_1, y_1)$, com $0 \leq y_1 < y_2 \leq 1$ e $0 \leq x_1 < x_2 \leq 1$ (ver figura 2). No interior do quadrado e em suas arestas e vértices, a função f é aproximada pela interpolação $\tilde{f}(x, y) = c_1x^2 + c_2y^2 + c_3x + c_4y + c_5$. Dados os valores da função f nos vértices do quadrado, 4 dos 5 coeficientes da interpolação podem ser encontrados; o quinto coeficiente vem da equação de autovalores, $\nabla^2 \tilde{f} = \lambda \tilde{f}$. Assim, para cada quadrado, tem-se um sistema de 5 variáveis e 5 equações em função dos valores de f nos 4 vértices, que na prática são desconhecidos. Conseqüentemente, exigindo a continuidade dos vértices entre quadrados compartilhados, tem-se um sistema de equações cuja solução é tanto o valor da função f em cada vértice quanto o conjunto de coeficientes de cada interpolação em cada célula.

Cuidados adicionais precisam ser tomados para garantir a continuidade de f ao longo de arestas adjacentes. Fixado o valor de y ou de x , a equação da interpolação $\tilde{f}(x, y) = c_1x^2 + c_2y^2 + c_3x + c_4y + c_5$ passa a ser um polinômio do segundo grau na variável não fixada. Por outro lado, um polinômio do segundo grau é completamente definido por três pontos distintos. Portanto, como nas arestas dos quadrados a continuidade dos vértices já é garantida, para garantir

a continuidade das arestas basta acrescentar mais um ponto, além de seus vértices, e exigir que este ponto assuma o mesmo valor nas interpolações dos quadrados adjacentes que o contém. Para cada quadrado, são 4 equações a mais, totalizando 9 equações e 9 variáveis (os valores de f nos 4 vértices e as constantes c_1, c_2, c_3, c_4 e c_5).

Novamente, encontra-se um problema de autovalores matricial, assim como no MDF. No entanto, neste caso o sistema de equações é usualmente maior, porque inclui também os coeficientes da interpolação válida para cada célula. As condições de contorno são tratadas da mesma forma que no MDF.

Existem diversas variações da técnica explicada nos últimos parágrafos. É possível, por exemplo:

- discretizar o domínio utilizando triângulos;
- utilizar um tipo diferente de interpolação, como por exemplo um polinômio de grau maior; e
- utilizar um sistema de coordenadas locais para cada célula.

A literatura de elementos finitos é vasta, mas a ideia básica apresentada aqui continua válida. Em última instância, o operador diferencial original é discretizado, gerando um problema de autovalores matricial que pode ser resolvido por algoritmos de álgebra linear [17].

Em comparação ao MDF, o MEF é mais complexo de ser implementado em um computador, pois demanda conhecimento de geometria computacional. Além disso, a montagem das matrizes geradas no MEF é mais complicada, pois tipicamente envolve células que trabalham com sistemas de coordenadas locais, muitas vezes não usuais, como é o caso dos elementos triangulares com coordenadas baricêntricas [19, 20].

Por outro lado, em aplicações que necessitem de uma função que possa ser avaliada em todos os pontos, o MEF é o mais indicado. Exemplos de aplicações deste tipo incluem aquelas que utilizam a função para estimar grandezas úteis por meio de integrais, como potência e energia. Outro exemplo são aplicações que usam as autofunções como base ortogonal para representar uma função qualquer da imagem do operador diferencial como combinação linear de autofunções. Neste caso, o cálculo dos coeficientes desta combinação também envolve integrais.

Atualmente, o MEF é o mais usado em aplicações comerciais, por ser bem conhecido e bem estudado, tanto por engenheiros quanto por matemáticos. Exemplos de aplicações incluem o COMSOL Multiphysics [21], o CST Microwave Studio [22] e o SolidWorks Simulation Premium [23]. Todavia, o método apresenta as mesmas limitações do MDF:

- necessidade de artifícios de truncamento caso o domínio seja infinito;
- o consumo de memória cresce com o aumento do número de autovalores necessários;
- alto consumo de memória, mesmo para estruturas simples.

Além disso, assim como no MDF, a solução encontrada depende de uma formulação precisa do problema. Caso algum parâmetro que define o problema seja desconhecido ou caso seja necessário fazer uma análise de sensibilidade para diversos valores de um parâmetro, é preciso repetir todo o processo: construir uma malha, gerar as matrizes, resolver o problema matricial de autovalores. Considere, por exemplo, que agora o operador de Laplace seja substituído por um outro operador muito semelhante,

$$\begin{aligned}\nabla_A^2 f &= f_{[xx]} + Af_{[yy]} \\ f &\in L^2([0; 1]^2) \\ f(0, y) &= f(1, y) = 0 \\ f(x, 0) &= f(x, 1) = 0 \\ x, y &\in [0; 1] \\ A &\in \mathbb{C}.\end{aligned}\tag{2-23}$$

O valor do parâmetro A poderia representar uma grandeza física, como um diâmetro, um comprimento ou uma constante física de um material. Caso a aplicação demande uma análise de sensibilidade para 100 valores de A , seria necessário executar o MEF 100 vezes. Para aplicações envolvendo vários parâmetros, facilmente o número de simulações poderia chegar a milhões, o que tipicamente inviabiliza o uso de técnicas como o MEF e o MDF.

Nesse sentido, esta tese busca abordar este problema utilizando uma abordagem que permita incluir parâmetros na solução, de maneira analítica. Tais soluções são conhecidas como *proxies* e são representações analíticas de funções parametrizadas, capazes de avaliar de maneira instantânea para qualquer combinação dos parâmetros. O método escolhido para obter essas proxies é o *Método da Programação Genética* (PG), cuja revisão bibliográfica está contemplada na próxima seção.

2.3

Emprego de Programação Genética (PG) na Literatura

Esta seção apresenta uma revisão bibliográfica sobre os usos particulares de PG na literatura que estão relacionados com o tema da presente pesquisa.

Não é objetivo desta seção apresentar o modelo de PG utilizado, o que será feito no próximo capítulo.

A técnica de PG [24] pertence ao grupo de técnicas conhecidas como computação evolucionária. Ela é geralmente classificada como uma heurística não-determinística para resolver problemas de otimização combinatória.

Neste trabalho, utiliza-se PG como parte de um modelo de solução de problemas de autovalores de operadores diferenciais vetoriais, conforme explicado na seção 2.1. No momento de escrita deste texto, não foi encontrado na literatura nenhum trabalho que aborde este exato problema utilizando um modelo baseado em PG. Contudo, existem alguns trabalhos a respeito de solução de equações diferenciais parciais com auxílio de PG [25, 26, 27]. Estes trabalhos diferem desta pesquisa nos seguintes aspectos:

- buscam resolver equações diferenciais escalares (funções de 1 dimensão), enquanto, neste trabalho, busca-se resolver problemas de autovalores vetoriais (funções de uma ou mais dimensões);
- não resolvem ou não são adaptáveis de maneira imediata para problemas de autovalores pois são concebidos para encontrar a única solução de uma equação diferencial parcial totalmente especificada;
- utilizam a técnica conhecida como diferenciação automática, que demanda diferenciar a função sempre que a derivada precise ser avaliada em algum ponto, o que torna o processo lento. Neste trabalho, utiliza-se um algoritmo de diferenciação simbólica, que gera a versão analítica da derivada e, por isso, há um ganho inerente de desempenho em relação às técnicas encontradas na literatura. Além disso, a versão analítica da derivada pode ser simplificada por meio de métodos simbólicos, o que remove singularidades e torna as funções mais simples de avaliar.

Nos próximos parágrafos, será feita uma análise dos principais trabalhos que se encaixam na descrição acima.

Em [25], o autor apresenta uma metodologia baseada em PG para resolver equações diferenciais parciais completamente especificadas. Ele usa diferenciação automática e um modelo baseado em estruturas de árvores para representar as funções analiticamente. Então, o autor utiliza amostragem da equação diferencial numa malha de pontos e calcula o erro de aproximação médio considerando cada ponto da malha. O mesmo ocorre para cada condição de contorno. Assim, o problema original transforma-se em um problema de minimização de erros no qual deseja-se encontrar a função que apresenta o menor erro em relação à solução exata da equação diferencial.

O trabalho [25] difere do presente trabalho por não abordar problemas de autovalores de fato, que têm infinitas soluções (tipicamente), ao contrário do problema de solução única abordado pelo autor de [25]. Além disso, [25] não aborda problemas vetoriais (funções de 1 ou mais dimensões). Por fim, a formulação da função objetivo (a função a ser minimizada) baseada em uma média de erros em uma malha de pontos não é adequada, porque demandaria evoluir os autovalores juntamente com as autofunções, o que aumentaria a complexidade do problema de otimização. Neste trabalho, o modelo evolui apenas as autofunções e calcula os autovalores por meio da razão de Rayleigh.

Além disso, destaca-se que o presente trabalho utiliza diferenciação simbólica ao invés de diferenciação automática. Isto ocorre porque, na diferenciação simbólica, o processo de diferenciação ocorre apenas uma vez, em oposição à diferenciação automática, que necessita diferenciar novamente sempre que a derivada precise ser avaliada para um determinado ponto.

Além do trabalho [25], outros trabalhos também aplicaram programação genética para resolver equações diferenciais parciais particulares. Em [26], os autores sugerem um algoritmo de PG também com diferenciação automática, semelhante a [25], para resolver equações de Hamilton-Jacobi-Bellman; Hamilton-Jacobi-Isaacs e Francis-Byrnes-Isidori em sistemas dinâmicos não lineares. O artigo conclui que a metodologia apresenta bons resultados e é uma boa alternativa para a solução de problemas de controle ótimo.

O artigo [27] também apresenta exemplos de soluções de equações diferenciais utilizando a metodologia de PG, mais uma vez com diferenciação automática. São abordados problemas de equações diferenciais escalares de primeira e segunda ordem. Diferentemente de [25, 26], a metodologia apresentada em [27] utiliza evolução gramatical para o algoritmo de PG, combinada com diferentes métodos de diferenciação automática. Os autores concluem que os resultados são bons, sendo iguais às soluções analíticas exatas em alguns casos.

Tendo em vista o contexto atual da literatura, a seção 2.4 apresenta um resumo contendo as contribuições originais deste trabalho.

2.4

Contribuições Originais

A presente seção resume este capítulo, evidenciando as contribuições originais. Este trabalho propõe um **modelo paramétrico de solução de problemas de autovalores vetoriais de operadores lineares diferenciais, utilizando PG**, conforme definido na seção 2. São características inovadoras deste modelo:

- o modelo resolve problemas de autovalores de operadores diferenciais vetoriais entregando uma solução analítica. As técnicas clássicas para solução deste tipo de problemas, mais especificamente o MDF e o MEF, não entregam solução analítica. Note que, embora o MEF entregue uma forma analítica em cada célula utilizando interpoladores, todo o processamento do algoritmo é numérico. No modelo aqui proposto, pretende-se adotar computação simbólica no processamento interno, delegando à computação numérica apenas a avaliação final da função.
- é o primeiro modelo para resolver problemas de autovalores de operadores diferenciais vetoriais na literatura baseado em programação genética e computação simbólica. Os modelos na literatura mais próximos não resolvem o problema de autovalores pois focam em resolver equações diferenciais. Além disso, estes modelos utilizam diferenciação automática, que é uma técnica numérica para encontrar a derivada, diferentemente no modelo proposto neste trabalho, que usa diferenciação simbólica em conjunto com simplificação simbólica para redução de singularidades.
- os modelos baseados em PG para resolver equações diferenciais convertem o problema original em um problema de regressão, no qual deseja-se minimizar o erro ponto a ponto. Para isso, o domínio das funções é discretizado numa malha de pontos, em geral uniforme. Neste trabalho, será adotada uma metodologia nova, sem discretização, que utiliza integrais e o conceito de distância entre funções. As integrais são estimadas de maneira adaptativa, utilizando a menor quantidade de pontos possível para garantir um erro máximo pré-especificado. Dessa forma, a metodologia proposta neste trabalho oferece garantias de erro de avaliação do indivíduo que não estão presentes nos modelos de solução de equações diferenciais da literatura.
- o modelo aqui proposto pode entregar como resposta funções parametrizadas por constantes do problema, como diâmetros, comprimentos e parâmetros físicos de materiais. Isso difere dos modelos baseados em MEF e MDF, que dependem de valores específicos dos parâmetros para construir a malha e as representações matriciais correspondentes. Como resultado, a resposta do modelo aqui proposto é uma função aproximada tanto dos pontos do domínio físico como dos parâmetros, o que permite ser usada em análises de sensibilidade e em problemas de otimização de maneira eficiente.
- o modelo proposto é aplicado em um estudo de caso real: a análise de propagação de ondas eletromagnéticas de poços de petróleo.

- o modelo aqui proposto também apresenta diversas inovações na área de computação evolucionárias que serão ressaltadas no capítulo referente à descrição do modelo. Dentre elas, está a representação de uma função vetorial por meio de indivíduos multi-árvore com compartilhamento de árvores.

Desta forma, considera-se que a presente pesquisa de doutorado apresenta contribuições importantes em relação à literatura referente à solução de problemas de autovalores de operadores diferenciais.

O próximo capítulo apresenta o modelo proposto nesta pesquisa, descrevendo também as inovações específicas da área de programação genética e computação evolucionária.

3

Modelo de Solução de Problemas de Autovalores

O propósito deste capítulo é responder à pergunta: **qual é o modelo proposto para resolver o problema em questão?** Sendo assim, este capítulo descreve em detalhes o modelo de solução de problemas de autovalores de operadores diferenciais lineares com auxílio de programação genética (PG). Mais especificamente, o capítulo enfatiza a descrição do modelo para operadores diferenciais lineares e normais. O capítulo também comenta uma possível adaptação do modelo para problemas envolvendo operadores que não são normais.

Como o modelo proposto é baseado em PG, opta-se por fazer uma breve introdução ao assunto antes de apresentar o modelo de fato. Em particular, tal introdução apresenta a heurística de PG como é tipicamente usada em problemas de regressão simbólica, pois é a base para o modelo proposto.

3.1

Introdução à PG Aplicada à Regressão Simbólica de Funções

3.1.1

Considerações Iniciais

A discussão desta seção baseia-se principalmente em [24] e foca em descrever a heurística de PG sob a ótica que se faz útil para esta tese. O objetivo é introduzir a nomenclatura e conceitos de PG no contexto de regressão simbólica por meio de indivíduos multi-árvore, para que o leitor possa se sentir familiarizado e entender os aspectos técnicos da contribuição da presente tese.

Programação Genética é uma técnica da área de Inteligência Computacional que engloba um conjunto de técnicas computacionais para resolver problemas de otimização combinatória. A inspiração para a técnica é a teoria da evolução de Charles Darwin [28], na qual uma população de indivíduos compete pelos recursos de um meio ambiente ao longo de gerações e apenas os mais aptos sobrevivem.

A PG é uma heurística estocástica de otimização global, particularmente útil para problemas de otimização combinatória. A PG é capaz de evoluir

programas de computador em linguagens de programação arbitrárias, de modo a mapear de maneira ótima entradas em saídas.

Em PG, submete-se uma população de indivíduos, que são representações de programas ou expressões matemáticas criadas aleatoriamente, a um processo artificial de seleção e evolução. Este processo iterativo divide-se em gerações, sendo cada uma subdividida nas seguintes etapas:

1. *seleção*: escolha de indivíduos promissores para procriação;
2. *recombinação e mutação*: formação de descendentes por meio de recombinação dos indivíduos selecionados e mutação;
3. *avaliação*: os novos indivíduos são avaliados utilizando uma *função de avaliação*, cujo objetivo é identificar o quão adaptado o indivíduo é em comparação aos outros indivíduos.
4. *substituição*: inserção dos novos indivíduos na população, em geral substituindo os indivíduos menos adaptados pelos mais adaptados.

Não existem garantias de obtenção de soluções ótimas ao longo das gerações, mas esta dinâmica tende a produzir respostas de boa qualidade.

Dentre os pontos fortes tipicamente citados da técnica de PG destacam-se:

- *paralelismo*: as etapas da heurística são paralelizáveis em vários níveis. Por exemplo, os indivíduos podem ser avaliados em paralelo na etapa de avaliação, que costuma ser a mais custosa.
- *produz soluções simbólicas*: os indivíduos são representados de maneira simbólica (códigos de computador e expressões matemáticas), logo as soluções obtidas pela PG são imediatamente legíveis e interpretáveis, além de serem menos suscetíveis a problemas numéricos.
- *é extensível e modificável*: a PG é muito versátil e, por isso, permite hibridizações (funcionamento junto a outras técnicas), e diversas representações e linguagens para os programas. *Uma das contribuições do modelo proposto nesta tese*, por exemplo, é utilizar uma combinação de técnicas de otimização local para estimar coeficientes do indivíduo durante a evolução.
- *é robusto*: como a heurística de PG não depende de uma sequência de iterações numéricas, ela é robusta a problemas típicos de computação numérica, como truncamentos, valores infinitos, e erros de arredondamento.

A seção 3.1.2 entra no tema “representação do indivíduo”, mostrando como a expressão analítica de uma função escalar é tipicamente codificada pela heurística de PG. Posteriormente, as demais seções mostram como funciona cada etapa que ocorre dentro de uma geração particular durante uma evolução por PG.

3.1.2 Representação do Indivíduo

Diversas representações para um indivíduo são possíveis, mas este trabalho se interessa apenas pela representação multi-árvore. Nesta representação, a expressão matemática do indivíduo é representada por uma combinação linear de árvores. Aqui o termo árvore é usado no sentido computacional, ou seja, cada a árvore é um conjunto de nós, representado de maneira recursiva, em que cada nó pode ter um número inteiro e finito de nós filhos.

Existem dois tipos de nós: os nós funcionais e os nós terminais. Nós terminais são aqueles que não têm nenhum nó filho e geralmente representam constantes do problema (exemplo π , gravidade etc.), e variáveis literais (ou seja, x , y , ρ etc.). Por outro lado, os nós funcionais têm um ou mais filhos, e representam operações matemáticas e funções elementares, como *sen*, *cos*, multiplicação, exponenciação, somatório, dentre outras.

Grande parte da flexibilidade da PG jaz nos conjuntos de nós funcionais e terminais. Se o usuário da técnica tiver algum conhecimento do problema a ser resolvido, pode ser interessante “ajudar” o otimizador por PG colocando na lista de possíveis nós funcionais e terminais os elementos que têm grande chance de participarem da solução. Grandes candidatos são funções que formem bases ortogonais conhecidas, como exponenciais e polinômios. Outras são funções que apareçam em soluções analíticas de problemas semelhantes, como funções de Bessel para problemas em coordenadas cilíndricas.

Como um exemplo ilustrativo, considere a representação da função $f(x) = 2x + \text{sen}(x)$ usando uma única árvore. Claramente, a função poderia ser reescrita como $f(x) = [2 * x] + [\text{sen}(x)]$ e dividida em forma de árvore conforme a figura 3.

Na representação da figura 3, as elipses representam nós funcionais, enquanto os nós sem elipse são os nós terminais.

Para conseguir representar de maneira simples neste texto, é comum utilizar a notação $f(x) = [(+ (* 2 x) (\text{sen } x))]$, semelhante à notação usada na linguagem de programação Lisp [29]. Note o uso de colchetes para delimitar uma árvore. No caso de múltiplas árvores, a notação utiliza 1 par de colchetes para cada árvore. Por exemplo, a função f também poderia ser escrita como

$f(x) = 2[x] + 1[(senx)]$, onde a árvore $[x]$ é formada apenas pelo terminal “ x ”.

3.1.3

Geração de uma População Inicial

A população inicial é a população utilizada na geração 0, e é criada de maneira aleatória. Existem alguns esquemas de geração das árvores, mas, basicamente, o processo funciona sorteando elementos de dois conjuntos: o de funções e o de terminais. Caso seja selecionado um elemento da lista de funções, o processo continua recursivamente até que se sorteie um terminal ou o nível máximo permitido para uma árvore seja alcançado. Este processo é repetido para cada árvore do indivíduo multi-árvores.

Os principais métodos de construção de um indivíduo árvore estão descritos a seguir.

- *Grow*: os nós são selecionados aleatoriamente dos conjuntos de funções e de terminais. Sendo assim, as árvores são de formato não regular. O processo de ramificação da árvore termina quando um símbolo terminal é encontrado ou o nível máximo de um ramo é atingido.
- *Full*: o método escolhe somente funções até que um nó de profundidade máxima seja selecionado, então ele passa a escolher somente terminais. Assim, todas as árvores têm o número máximo de nós e a profundidade máxima. Recomenda-se, contudo, que as árvores iniciais sejam pequenas.

Além dos esquemas mostrados, existem outros esquemas de inicialização mais sofisticados. Por exemplo, caso seja interessante gerar indivíduos de maneira equiprovável, o método uniforme é interessante [30], apesar de seu alto custo computacional. Neste trabalho, usa-se, geralmente, o método *full*.

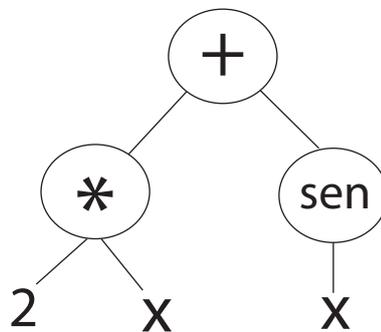


Figura 3: Representação com uma árvore para a função $f(x) = 2x + \text{sen}(x)$.

3.1.4

Operador de Recombinação

O operador de recombinação usado neste trabalho é conhecido como “recombinação por sub-árvores”. Neste tipo de operador, dois indivíduos são cruzados, gerando dois filhos. As árvores de mesma posição no cromossomo são recombinadas, gerando duas novas árvores, cada uma delas pertencente a um indivíduo filho distinto.

Para recombinar duas árvores, escolhe-se dois pontos de cortes aleatórios, sendo um deles de cada árvore. Posteriormente, as árvores são cortadas nos pontos de corte e as sub-árvores são invertidas ou trocadas entre as árvores pais. Assim, após a operação, os filhos serão formados por material genético das duas árvores pai. Para maiores detalhes, vide [24].

3.1.5

Operadores de Mutação

Os operadores de mutação tipicamente adotados na literatura são 6 e estão descritos em [31]. São eles:

- *OneNode*: aleatoriamente seleciona um nó de alguma árvore do indivíduo e substitui por outro nó. Se o nó for terminal, substitui por terminal. Se for uma função de n entradas, substitui por outra função de n entradas.
- *AllNodes*: igual ao *onenode*, mas aplica o processo em todos os nós, ao invés de um único nó.
- *Swap*: seleciona aleatoriamente um nó funcional com 2 ou mais entradas (filhos) e permuta aleatoriamente os filhos.
- *Grow*: seleciona aleatoriamente um terminal e o substitui por uma árvore aleatória. Efetivamente, este processo aumenta o número de nós da árvore.
- *Trunc*: o oposto de *grow*. Este método seleciona aleatoriamente um nó funcional e o substitui por um terminal, reduzindo efetivamente o tamanho da árvore.
- *Gaussian*: seleciona aleatoriamente um terminal numérico e o substitui por um número sorteado aleatoriamente de uma distribuição normal centrada no valor original do nó.

Dadas os tipos de mutação, é importante ressaltar que nem sempre as mutações são aplicadas. Na realidade, para cada indivíduo, de 0 à 6 mutações podem ser aplicadas, com maior probabilidade para quantidades menores de

mutação. Em geral, a probabilidade de alguma mutação ser aplicada é pequena (menor do que 10%).

Outro ponto importante é que, à medida que a população evolui e se aproxima do ótimo global, é comum que sua diversidade diminua. Tal característica não é boa, pois se todos os indivíduos forem idênticos, então o algoritmo perde a capacidade de evoluir. Por essa razão, é comum adotar uma política adaptativa na taxa de mutação, que permita que a taxa de mutação seja relativamente alta em cenários de pouca diversidade e relativamente baixa em cenários com muita diversidade. Idealmente, a mutação deve ocorrer raramente e a diversidade da população deve ser alta.

3.1.6

Seleção de Indivíduos

A seleção de indivíduos é o processo pelo qual dois pais são escolhidos aleatoriamente na população para gerar dois filhos que serão introduzidos na mesma. Em geral, para uma população de N indivíduos, seleciona-se $N/2$ pares de indivíduos para procriação, gerando, portanto, N novos indivíduos que irão competir com os outros N já existentes por N posições na população.

A seleção de indivíduos geralmente é projetada para escolher os mais aptos (melhor avaliação) como pais. Como indivíduos menos aptos podem ainda conter material genético interessante, pode ser útil selecioná-los como pais, mas com menor probabilidade.

Um dos métodos de seleção mais usados é o método do torneio, no qual os N pais são escolhidos sequencialmente, a partir de N torneios distintos. Cada torneio tem um número pequeno de participantes (tipicamente de 2 a 10), que podem ser escolhidos aleatoriamente da população ou a partir de alguma heurística. Os participantes do torneio são, então, ranqueados do melhor para o pior, baseado na aptidão de cada um e em outros critérios importantes, como o tamanho do indivíduo (indivíduos menores são em geral melhores). Finalmente, o torneio escolhe seu vencedor por meio de um sorteio simples, no qual indivíduos com posição mais alta no ranking tem maior chance de vencer do que indivíduos com posição mais baixa.

Outros métodos são possíveis, para uma lista mais completa, veja [24].

3.1.7

Avaliação

A avaliação é o processo que analisa a aptidão de cada indivíduo, associando à ele um número real. Em problemas de minimização, deseja-se

minimizar este número. É a parte mais configurável de um algoritmo evolutivo e é, em geral, a parte de maior custo computacional.

Em problemas de regressão simbólica, a avaliação retorna o erro de regressão do indivíduo. Este erro pode ser calculado, por exemplo, pelo erro quadrático médio. Trabalhos neste sentido incluem [24, 32, 33, 25]. Nesta abordagem, é necessário uma tabela de valores de referência, formada por muitos pares $(x, f(x))$. Assim, escolhe-se o indivíduo que represente a f com o menor erro possível.

Algumas adaptações desta ideia incluem o trabalho [25], que utilizou um modelo de regressão para resolver equações diferenciais parciais escalares. No referido trabalho, o autor construiu artificialmente a tabela de valores $(x, f(x))$ a partir da equação diferencial e buscou a função que resolvia a equação com menor erro de regressão possível.

No caso de indivíduos multi-árvore, a função f é uma combinação linear das expressões representadas por cada uma de suas árvores. Como os coeficientes desta combinação não fazem parte da representação do indivíduo, geralmente eles são encontrados por meio de minimização local durante a avaliação do indivíduo, ou seja, escolhe-se os valores dos coeficientes de modo a minimizar o erro de regressão, dadas as árvores. Neste caso, a avaliação é o erro mínimo obtido por meio desta otimização. Obviamente, este processo é repetido para cada indivíduo a avaliar, o que naturalmente torna a avaliação um processo lento, como já havia sido ressaltado. Contudo, note que todos os indivíduos podem ser avaliados em paralelo.

3.1.8 Substituição

O processo de substituição é o último que ocorre em uma geração da heurística de programação genética. A finalidade da substituição é escolher quais indivíduos, dentre os novos e os antigos, devem ficar na população.

Para fazer essa escolha, o método de substituição deve priorizar os indivíduos mais aptos, independentemente de serem novos indivíduos ou não. Por isso, uma forma comum de fazer a substituição é simplesmente escolher os melhores N indivíduos dentre os $2N$ existentes, onde N é o tamanho da população. No entanto, formas mais sofisticadas podem existir com a finalidade de balancear a escolha dos indivíduos mais aptos com a diversidade da população resultante. Um exemplo é a substituição com *steady state*, que garante que os k melhores indivíduos da última geração não serão substituídos, mesmo que existam mais do que $N - k$ indivíduos melhores.

Substituição é um tópico de pesquisa ativo na área de computação evolucionária. Para maiores informações, vide [24].

3.2

Descrição do Modelo de Solução de Problemas de Autovalores

Nesta seção, este trabalho apresenta o *modelo paramétrico de solução de problemas de autovalores de operadores diferenciais vetoriais assistido por programação genética*. Para isso, apresento primeiro o modelo desenvolvido para operadores auto-adjuntos não-negativos. Este caso também engloba os operadores não-positivos, uma vez que, se A é um operador linear não-negativo, então $-A$ é um operador linear não-positivo. Posteriormente, o capítulo apresenta uma discussão para mostrar como resolver para operadores auto-adjuntos indefinidos. Então, é apresentado o modelo de solução para operadores normais, baseado no modelo para operadores auto-adjuntos. Finalmente, o capítulo termina mostrando algumas ideias de como adaptar o modelo para outros tipos de operadores lineares e diferenciais.

Antes de continuar, é preciso enfatizar que *todo o conteúdo desta seção é contribuição original deste trabalho*.

3.2.1

Modelo Proposto assumindo Operadores Auto-Adjuntos Não-Negativos

3.2.1.1

Considerações Iniciais

Nesta seção, queremos resolver o problema de autovalores padrão $Af = \lambda f$, onde A é um operador linear auto-ajunto não-negativo definido em um espaço de Hilbert H com produto interno $\langle \cdot; \cdot \rangle$. Deseja-se encontrar os autopares $(\lambda; f)$, que tipicamente formam um conjunto discreto e infinito.

Além disso, assume-se que o espaço de Hilbert H no qual o operador A é definido é o espaço $L^2(\Omega)^N$, formado por funções de energia finita ($\langle f; f \rangle < \infty$), definidas no domínio $\Omega \subseteq \mathbb{R}^M$ e com contradomínio \mathbb{C}^N . Neste trabalho, tais funções são chamadas de *funções vetoriais*. Além disso, exige-se que as funções f sejam *duas vezes diferenciáveis* e a *primeira derivada seja contínua* em qualquer direção.

Todo o modelo de solução é construído em cima de uma propriedade muito importante já discutida no capítulo 2: em operadores não-negativos, o mínimo da razão de Rayleigh é igual ao menor autovalor. Portanto, dado o

funcional

$$F_A[f] = \frac{\langle Af; f \rangle}{\langle f; f \rangle},$$

$$f \in L^2(\Omega)^N, \quad (3-1)$$

o problema de encontrar a função f que o minimiza é equivalente ao problema de encontrar o menor autovalor de A .

Tendo o menor autovalor de A , uma modificação particular pode ser aplicada de modo a remover o autovetor associado a este autovalor do espectro de A , gerando um operador novo, A' . Posteriormente, o processo se repete: o método busca minimizar o funcional 3-1, utilizando agora o novo operador A' no lugar de A , assim obtendo o segundo menor autovalor. O processo é repetido até que uma quantidade suficiente de autopares tenha sido obtida.

Sendo assim, o problema principal é encontrar uma função f que minimize o funcional 3-1. Esse problema é resolvido no modelo proposto de maneira simbólica, utilizando uma adaptação do modelo de PG geralmente usado em problemas de regressão simbólica.

No contexto de operadores diferenciais, tipicamente um conjunto de condições de contorno faz parte da definição do domínio do operador. Em outras palavras, deseja-se que a função f também satisfaça uma série de restrições. Por exemplo, condições de contorno de Dirichlet especificam valores constantes que a função deve ter nos limites de seu domínio. Condições de contorno de Neumann especificam valores constantes para as derivadas nos limites do domínio. Para tratar estas condições, proponho considerar as condições de contorno como restrições ao problema de otimização.

Sendo assim, deseja-se minimizar o valor da razão de Rayleigh sujeito à restrições que, quando atendidas, implicam em soluções que respeitem as condições de contorno do problema. A seguir, os próximos parágrafos irão formalizar o modelo proposto para resolver este problema.

3.2.1.2 Modelo de PG

Considere que as condições de contorno do problema podem ser representadas por uma equação do tipo $Bf = 0$, onde B é um segundo operador linear, não-necessariamente auto-adjunto, mas definido também no espaço $L^2(\Omega)^N$.

Queremos encontrar a solução do problema de otimização

$$\begin{aligned}
 & \mathbf{min} \ F_A[f] \\
 & \text{sujeito à} \\
 & \quad Bf = 0 \\
 & \quad f \in L^2(\Omega)^N.
 \end{aligned} \tag{3-2}$$

Para isso, a função vetorial $f = [f_1 \ f_2 \ f_3 \ \dots \ f_N]^T$ é codificada em um modelo multi-árvore contendo k_a árvores por função escalar, $k_a \in \mathbb{Z}$. É possível que haja sobreposição de k_s árvores, ou seja, duas funções escalares f_j e f_{j+1} compartilham k_s árvores, $k_s \in \mathbb{Z}$, $0 \leq k_s \leq k_a$. Obviamente, se $k_s = 0$, então não há sobreposição e se $k_s = k_a$, então $f_1 = f_2 = \dots = f_N$. Então, se h_m é a função representada pela m -ésima árvore, temos:

$$\begin{aligned}
 f_1 &= d_{1,1} h_1 + d_{1,2} h_2 + \dots + d_{1,k_a} h_{k_a} \\
 f_2 &= d_{2,1} h_{(k_a-k_s+1)} + d_{2,2} h_{(k_a-k_s+2)} + \dots + d_{2,k_a} h_{(2k_a-k_s)} \\
 &\vdots \\
 f_N &= d_{N,1} h_{[(N-1)k_a-(N-1)k_s+1]} + \dots + d_{N,k_a} h_{[Nk_a-(N-1)k_s]} \\
 d_{l,m} &\in \mathbb{C} \\
 l &= 1, \dots, N \quad m = 1, \dots, k_a.
 \end{aligned} \tag{3-3}$$

A figura 4 apresenta um exemplo de cromossomo com $k_a = 3$ e $k_s = 1$.

Os coeficientes complexos $d_{l,m}$ não fazem parte da codificação do indivíduo. Assim como no modelo multi-árvore para regressão simbólica já discutido neste capítulo, estes coeficientes serão obtidos usando uma segunda otimização, neste caso numérica, utilizando um algoritmo baseado em gradientes. Além disso, os valores de k_a e k_s são pré-especificados, ou seja, o número de árvores é fixo durante a evolução via PG. Essa é uma *importante diferença para modelos multi-árvores utilizados na literatura* para fazer regressão simbólica de funções escalares, pois estes utilizam uma quantidade de árvores que também faz parte da otimização.

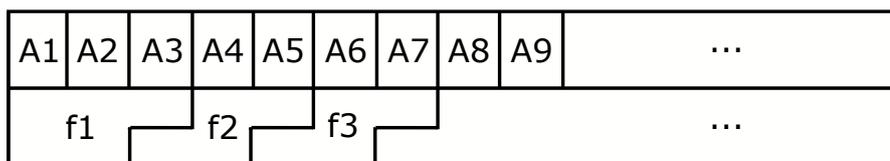


Figura 4: Exemplo de cromossomo com $k_a = 3$ e $k_s = 1$. A função escalar f_1 é combinação linear das árvores A_1, A_2, A_3 , enquanto que f_2 é combinação linear de A_3, A_4, A_5 e f_3 é combinação linear de A_5, A_6 e A_7 .

Algoritmo 1 Avaliação do indivíduo multi-árvore vetorial

-
- 1: **procedure** AVALIAR(*ind*) **retorna** $aval \in \mathbb{R}$.
 - 2: *ind* tem as árvores (vide 3-3)
 - 3: $f \leftarrow$ combinação linear que minimiza $F_A[f]$ sujeito à $Bf = 0$.
 - 4: $aval \leftarrow F_A[f]$.
 - 5: **return** $aval$
-

Dada a representação do indivíduo, a função de avaliação consiste em calcular o valor do funcional $F_A[f]$, onde F_A é o funcional definido em 3-1, para a melhor combinação linear obtida escolhendo os coeficientes de 3-3. A melhor combinação de coeficientes é aquela que minimiza $F_A[f]$ dadas as árvores sujeito à $Bf = 0$. O algoritmo 1 mostra abstratamente este processo.

Portanto, **o indivíduo é formado apenas pelas árvores**, enquanto que os coeficientes $d_{l,m}$ são otimizados durante a execução da função de avaliação de modo que a função vetorial f , definida como em 3-3, minimize o funcional $F_A[f]$. O resultado da avaliação do indivíduo é, então, o valor mínimo de $F_A[f]$. **É importante repetir** que, enquanto a PG é responsável por modificar a expressão analítica das árvores por meio de recombinação e mutação, um algoritmo de otimização local é o responsável por montar a função de fato, utilizando as árvores fornecidas e escolhendo os coeficientes ótimos.

Para fazer a otimização local, sugiro utilizar qualquer algoritmo baseado em gradiente para problemas de otimização sem restrições. Um exemplo é o algoritmo *L-BFGS-B Quasi-Newton* [34, 35], capaz de aproximar o gradiente e a matriz Hessiana eficientemente durante sua execução. Outro exemplo é o método *Trust Region* [36].

Para avaliar os operadores A e B aplicados à função f , utiliza-se diferenciação simbólica das árvores. Para isso, dada uma árvore, o presente modelo aplica a regra da cadeia recursivamente na árvore, levando em conta as derivadas das funções conhecidas e a variável de derivação. Caso a 2a. derivada seja necessária, o processo pode ser repetido sobre a árvore gerada após a aplicação da primeira derivada. Com isso, tem-se simbolicamente as funções Af e Bf , como função dos coeficientes da expansão 3-3. Além da diferenciação simbólica, o presente modelo também aplica simplificação simbólica nas árvores diferenciadas para remover possíveis singularidades, como divisão por 0 em expressões como x/x .

Durante a execução do algoritmo de minimização local, é necessário avaliar o valor numérico de Af e Bf a partir de suas representações simbólicas e dos valores numéricos dos coeficientes $d_{l,m}$. Esse processo é feito utilizando um algoritmo adaptativo de integração numérica para calcular os produtos internos presentes em $F_A[f]$. Este algoritmo subdivide o domínio de integração

recursivamente, aplicando uma quadratura para estimar a integral de cada subdomínio com limite superior de erro. O processo termina quando o erro global é menor do que uma constante, tipicamente da ordem de 10^{-5} . Neste trabalho, proponho o uso da quadratura de Gauss-Kronrod [37].

As operações genéticas utilizados pelo modelo proposto são aquelas apresentadas anteriormente neste capítulo. A recombinação é realizada por meio de cruzamento sub-árvore [24]. O cruzamento é feito para cada par de árvores na mesma posição dos indivíduos pais, gerando duas novas árvores na mesma posição para os dois filhos. Assim, as árvores na j -ésima posição dos indivíduos pais são cruzadas para gerar as árvores na j -ésima posição de seus dois filhos. Este cruzamento foi explicado na seção 3.1.

Para realizar a mutação, emprega-se os 6 operadores descritos na seção 3.1. Eles são aplicados de forma independente em cada árvore do indivíduo, ou seja, pode acontecer (inclusive, é frequente) de apenas algumas árvores do indivíduo sofrerem mutação. Além disso, a mutação ocorre sempre após a recombinação e tem uma taxa ou probabilidade de ocorrência pré-especificada que pode variar de maneira adaptativa para manter alta a diversidade da população sem prejudicar a evolução.

O processo de seleção de indivíduos para recombinação segue o esquema de torneio, conforme explicado na seção 3.1. O torneio é montado de modo a priorizar os indivíduos com a menor avaliação e o menor número de nós (árvores menores são mais desejáveis).

Para finalizar, o algoritmo 2 mostra de maneira abstrata como funciona a heurística de busca do indivíduo autofunção de A sujeito à $Bf = 0$ baseada em PG. Cada etapa descrita neste algoritmo fora detalhada nos parágrafos anteriores. Note que, por ser uma heurística, este trabalho não demonstra seu funcionamento com rigor matemático. Contudo, é possível demonstrar empiricamente que a heurística funciona em grande parte dos casos práticos, o que será feito nos próximos capítulos. No entanto, antes disso este texto precisa mostrar como calcular outros autovalores além do primeiro.

3.2.1.3 Cálculo dos Demais Autovalores

Dada a autofunção referente ao menor autovalor, é possível modificar o operador A de modo que esta não faça parte de seu conjunto de autofunções. Para isso, deve-se garantir que, se f_1 é a autofunção referente ao menor autovalor, então todas as funções do tipo af_1 , $a \in \mathbb{C}$, não fazem mais parte do domínio de A . No caso mais geral, se f_1, f_2, \dots, f_q são q autofunções linearmente independentes referentes aos q menores autovalores, então o domínio de A não pode conter nenhuma combinação linear das referidas autofunções.

Para resolver este problema, proponho adotar um operador de projeção, capaz de gerar funções ortogonais a um conjunto de funções linearmente independentes conhecidas. Na prática, considere o operador P_u definido no mesmo espaço de Hilbert do operador A como

$$P_u f = f - \frac{\langle f; u \rangle}{\langle u; u \rangle} u$$

$$f, u \in D(A), \quad (3-4)$$

onde $D(A)$ é o domínio de A , conforme especificado anteriormente. O operador P_u é o operador de Gram-Schmidt [5, 4, 6], utilizado no processo de ortogonalização de mesmo nome. Uma propriedade fundamental deste operador é que $P_u f$ e u são ortogonais, ou seja, $\langle P_u f; f \rangle = 0$. Em outras palavras, o conjunto imagem de P_u é formado apenas por funções ortogonais a u e, por isso, linearmente independentes a u .

Conseqüentemente, se f_1 é a autofunção relacionada ao menor autovalor, então, para toda função $f \in D(A)$, $P_{f_1} f$ é ortogonal e linearmente independente à autofunção f_1 . Portanto, ao invés de procurar a segunda autofunção

Algoritmo 2 Heurística de minimização por PG

- 1: **procedure** EVOLUIR(N_{pop}, p_{mut}) **retorna** $melhor \in L^2(\Omega)^N$.
 - 2: $pop \leftarrow$ população aleatória com N_{pop} indivíduos.
 - 3: $ger \leftarrow 0$
 - 4: **repeat**
 - 5: $pais \leftarrow$ seleção de $N_{pop}/2$ pais para recombinação.
 - 6: $filhos \leftarrow$ recombinação dos pares de $pais$.
 - 7: Aplica-se mutação nos $filhos$ com probabilidade p_{mut} .
 - 8: $avals \leftarrow$ Avaliação dos $filhos$.
 - 9: Aplica-se substituição em pop para incluir os novos $filhos$.
 - 10: $ger \leftarrow ger + 1$.
 - 11: **until** $ger > max$
 - 12: $melhor \leftarrow$ indivíduo válido de pop com a melhor avaliação.
 - 13: **return** $melhor$
-

no domínio definido para o operador A , procura-se no conjunto imagem da projeção P_{f_1} , $R(P_{f_1})$. Em outras palavras, procura-se resolver o problema de autovalores generalizado

$$\begin{aligned} A'f &= \lambda P_{f_1}f \\ B'f &= 0 \\ \text{com} \\ A' &= AP_{f_1} \\ B' &= BP_{f_1} \\ f &\in L^2(\Omega)^N. \end{aligned} \quad (3-5)$$

Note que o domínio especificado no problema de autovalores 3-5 é idêntico ao domínio do problema de autovalores original. Na prática, para resolver este problema, busca-se uma função f no domínio do operador original A cuja projeção $P_{f_1}f$ é a autofunção de A de menor autovalor.

Para encontrar o terceiro menor autovalor e a autofunção associada, basta repetir o processo, definindo o operador $A'' = A'P_{f_2}$. Note que

$$\begin{aligned} A''f &= AP_{f_1}P_{f_2}f \\ &= AP_{f_1} \left(f - \frac{\langle f; f_2 \rangle}{\langle f_2; f_2 \rangle} f_2 \right) \\ &= A \left(f - \frac{\langle f; f_1 \rangle}{\langle f_1; f_1 \rangle} f_1 - \frac{\langle f; f_2 \rangle}{\langle f_2; f_2 \rangle} f_2 + \frac{\langle \frac{\langle f; f_2 \rangle}{\langle f_2; f_2 \rangle} f_2; f_1 \rangle}{\langle f_1; f_1 \rangle} f_1 \right) \\ &= A \left(f - \frac{\langle f; f_1 \rangle}{\langle f_1; f_1 \rangle} f_1 - \frac{\langle f; f_2 \rangle}{\langle f_2; f_2 \rangle} f_2 + \frac{\langle f; f_2 \rangle}{\langle f_2; f_2 \rangle} \frac{\langle f_2; f_1 \rangle}{\langle f_1; f_1 \rangle} f_1 \right) \\ &= A \left(f - \frac{\langle f; f_1 \rangle}{\langle f_1; f_1 \rangle} f_1 - \frac{\langle f; f_2 \rangle}{\langle f_2; f_2 \rangle} f_2 \right) \text{ pois } \langle f_2; f_1 \rangle = 0. \end{aligned} \quad (3-6)$$

Como

$$f - \frac{\langle f; f_1 \rangle}{\langle f_1; f_1 \rangle} f_1 - \frac{\langle f; f_2 \rangle}{\langle f_2; f_2 \rangle} f_2 \quad (3-7)$$

é exatamente a expressão para o terceiro vetor no processo de ortogonalização de Gram-Schmidt, estamos efetivamente fazendo um processo de ortogonalização nas autofunções enquanto procuramos por elas. Assim, o modelo aqui proposto não só é capaz de encontrar uma quantidade arbitrária de autopares como também os encontra em ordem crescente de autovalores e garante que as autofunções formem um conjunto ortogonal de funções.

Para finalizar esta seção, revisito o algoritmo de avaliação do indivíduo,

agora considerando a projeção combinada $\tilde{P}_l = P_{f_1} P_{f_2} \dots P_{f_l}$ para $l = 1, 2, \dots$, ou seja, busca-se uma autofunção ortogonal às l autofunções conhecidas. O algoritmo 3 está descrito abaixo.

É importante ressaltar que, após a execução do algoritmo, a autofunção desejada será $\tilde{P}_l f$ ao invés de f , pois agora estamos procurando f de modo que $\tilde{P}_l f$ seja uma autofunção. Com isso, encerra-se a descrição do modelo de solução de problemas de autovalores de operadores diferenciais auto-adjuntos não-negativos. A próxima seção irá mostrar como adaptar o modelo para operadores não-positivos.

3.2.2

Modelo Proposto assumindo Operadores Auto-Adjuntos Não-Positivos

No caso de operadores não-positivos, a proposta é converter o problema de encontrar as soluções de $Af = \lambda f$ no problema de encontrar as soluções de $(-A)f = (-\lambda)f$. Neste caso, o operador $-A$ é não-negativo e todo o processo descrito anteriormente pode ser aplicado, com o cuidado que os autovalores de $-A$ devem ser multiplicados por -1 para obter os autovalores de A .

3.2.3

Modelo Proposto assumindo Operadores Auto-Adjuntos Indefinidos

Para operadores auto-adjuntos indefinidos, não existe, necessariamente, o menor (ou maior) autovalor. Por essa razão, não faz mais sentido encontrar o mínimo da *razão de Rayleigh*, uma vez que este mínimo pode sequer existir. Sendo assim, a metodologia apresentada para operadores auto-adjuntos não-negativos (e não-positivos) é inadequada e precisa ser modificada.

Uma alternativa seria buscar minimizar o módulo da *razão de Rayleigh*. Neste caso, intuitivamente estaríamos buscando o autovalor mais próximo de 0. Contudo, esta abordagem não necessariamente funciona, porque não há garantias matemáticas que a função que minimiza o módulo do funcional 3-1 seja uma autofunção.

Sendo assim, este trabalho propõe uma outra abordagem, menos trivial, que envolve o operador composto AA . Considere o operador auto-adjunto

Algoritmo 3 Avaliação do indivíduo multi-árvore vetorial, dadas as autofunções f_1, f_2, \dots, f_l

- 1: **procedure** AVALIAR(*ind*) **retorna** $aval \in \mathbb{R}$.
 - 2: *ind* tem as árvores (vide 3-3)
 - 3: $f \leftarrow$ combinação de árvores que minimiza $F_A[\tilde{P}_l f]$ sujeito à $B\tilde{P}_l f = 0$.
 - 4: $aval \leftarrow F_A[\tilde{P}_l f]$.
 - 5: **return** $aval$
-

indefinido A , com domínio e contradomínio iguais ao espaço de Hilbert H e com produto interno $\langle \cdot; \cdot \rangle$. O operador AA , formado pela composição do operador A com ele mesmo, é definido para cada $x \in H$ como $(AA)x = A(Ax)$. Ou seja, para calcular AAx deve-se calcular Ax e, então, aplicar novamente o operador A ao resultado, obtendo AAx .

O operador composto AAx tem uma série de propriedades úteis, as quais serão demonstradas a seguir.

Corolário 3.1 (AA é auto-adjunto) *Seja H um espaço de Hilbert e A um operador auto-adjunto, com domínio e contradomínio iguais à H . Então, o operador composto AA , definido para todo $x \in H$ por $AAx = A(Ax)$, é auto-adjunto.*

Prova. Seja $\langle \cdot; \cdot \rangle$ o produto interno de H . Como A é auto-adjunto, por definição temos

$$\begin{aligned} \langle Ax; y \rangle &= \langle x; Ay \rangle \\ x, y &\in H. \end{aligned} \tag{3-8}$$

Logo

$$\begin{aligned} \langle AAx; y \rangle &= \langle Ax; Ay \rangle && \text{pois } A \text{ é auto-adjunto} \\ &= \langle x; AAy \rangle && \text{pois } A \text{ é auto-adjunto,} \end{aligned} \tag{3-9}$$

ou seja, para cada $x, y \in H$, temos $\langle AAx; y \rangle = \langle x; AAy \rangle$, o que prova que AA é auto-adjunto. ■

Corolário 3.2 (AA é não-negativo) *Seja H um espaço de Hilbert e A um operador auto-adjunto, com domínio e contradomínio iguais à H . Então, o operador composto AA , definido para todo $x \in H$ por $AAx = A(Ax)$, é não-negativo.*

Prova. Seja $\langle \cdot; \cdot \rangle$ o produto interno de H . Logo, usando a prova do corolário 3.1, temos

$$\begin{aligned} \langle AAx; x \rangle &= \langle Ax; Ax \rangle && \text{pois } A \text{ é auto-adjunto} \\ &= \|Ax\|^2 && \text{norma induzida} \\ &\geq 0, \end{aligned} \tag{3-10}$$

ou seja, para cada $x \in H$, temos $\langle AAx; x \rangle \geq 0$, o que prova que AA é não-negativo. ■

Portanto, os corolários 3.1 e 3.2 mostram que o operador AA é auto-adjunto e não-negativo e, assim, a metodologia apresentada pode ser utilizada para determinar seus autovalores. Resta então verificar se há uma relação direta entre os autopares de AA e os autopares de A , o que será mostrado no próximo corolário.

Corolário 3.3 (autopares de AA) *Seja H um espaço de Hilbert e A um operador auto-adjunto, com domínio e contradomínio iguais à H e considere o operador composto AA , definido para todo $x \in H$ por $AAx = A(Ax)$. Então, para cada $\lambda \in \mathbb{R}$ e $x \in H$, se (λ, x) é um autopar de A , então (λ^2, x) é um autopar de AA .*

Prova. Se (λ, x) é um autopar de A , então, por definição, $Ax = \lambda x$. Sendo assim, $AAx = \lambda Ax = \lambda^2 x$, ou seja, (λ^2, x) é um autopar de AA . ■

Com o resultado do corolário 3.3, concluímos que toda autofunção de A também é uma autofunção de AA . Portanto, podemos buscar as autofunções do operador auto-adjunto não-negativo AA e “testar” se estas são autofunções de A . Para verificar se uma função $h \in H$ é autofunção de A , basta verificar se

$$\|Ah - F_A[h]h\| \leq \varepsilon, \quad (3-11)$$

onde ε é uma constante positiva pré-estabelecida próxima de 0 (tipicamente 10^{-8}) e $F_A[\cdot]$ fora definido na eq. 2-8.

Com isso, o modelo de determinação de autopares de operadores auto-adjuntos utilizando PG está totalmente definido. A seguir, será abordado o caso mais geral de operadores normais.

3.2.4

Modelo Proposto assumindo Operadores Normais

Operadores normais, conforme explicado na seção 2.1.1, são aqueles que comutam com o seu adjunto. Isto é, se A é um operador normal com domínio e contradomínio iguais ao espaço de Hilbert H e A^* é o seu adjunto, então, para cada $x \in H$, $AA^*x = A^*Ax$. Evidentemente, todo operador auto-adjunto é normal, pois neste caso $A^* = A$. Como a metodologia para operadores auto-adjuntos já foi explicada anteriormente, esta seção irá se preocupar apenas com aqueles operadores normais que não são auto-adjuntos.

No caso de operadores normais em geral, não há garantias matemáticas que o conjunto de autovalores seja real, ou seja, é possível que alguns autovalores sejam números complexos. Assim, não faz mais sentido minimizar a *razão de Rayleigh* para encontrar o menor autovalor. No entanto, se o operador adjunto for conhecido, é possível modificar o operador para torna-lo auto-adjunto e não-negativo, como mostra o próximo corolário.

Corolário 3.4 (operador AA^*) *Sejam H um espaço de Hilbert, A um operador normal com domínio e contradomínio iguais a H e A^* o seu adjunto. Então, o operador composto AA^* , definido para cada $x \in H$ por $AA^*x = A(A^*x)$, é auto-adjunto e não-negativo.*

Prova. Por definição de operador adjunto,

$$\begin{aligned} \langle Ax; y \rangle &= \langle x; A^*y \rangle \\ x, y &\in H. \end{aligned} \tag{3-12}$$

Logo, para cada $x, y \in H$,

$$\begin{aligned} \langle AA^*x; y \rangle &= \langle A^*x; A^*y \rangle && \text{definição de adjunto} \\ &= \langle x; AA^*y \rangle && \text{definição de adjunto,} \end{aligned} \tag{3-13}$$

ou seja, para cada $x, y \in H$ temos $\langle AA^*x; y \rangle = \langle x; AA^*y \rangle$, então AA^* é auto-adjunto. Além disso,

$$\begin{aligned} \langle AA^*x; x \rangle &= \langle A^*x; A^*x \rangle && \text{definição de adjunto} \\ &= \|A^*x\|^2 && \text{norma induzida} \\ &\geq 0, \end{aligned} \tag{3-14}$$

isto é, para cada $x \in H$ temos $\langle AA^*x; x \rangle \geq 0$, logo AA^* é não-negativo. ■

Teorema 3.5 (espectro de AA^*) *Sejam H um espaço de Hilbert, A um operador normal com domínio e contradomínio iguais a H e A^* o seu adjunto. Logo, se (λ, x) é um autopar de A , então $(|\lambda|^2, x)$ é um autopar de AA^* .*

Prova. Seja (λ, x) um autopar de A e seja N um operador com domínio e contradomínio iguais à H definido para cada $y \in H$ por $Ny = A^*(A - \lambda I)y$, onde I é o operador identidade ($Iy = y, y \in H$). Logo, o adjunto de N , N^* , é dado para cada $y \in H$ por $N^*y = (A^* - \bar{\lambda}I)Ay$ e, além disso,

$$\begin{aligned} NN^*y &= A^*(A - \lambda I)(A^* - \bar{\lambda}I)Ay \\ &= (A^*AA^*A - \lambda A^*A^*A + \lambda\bar{\lambda}A^*A - \bar{\lambda}A^*AA)y \\ N^*Ny &= (A^* - \bar{\lambda}I)AA^*(A - \lambda I)y \\ &= (A^*AA^*A - \lambda A^*AA^* + \lambda\bar{\lambda}AA^* - \bar{\lambda}AA^*A)y. \end{aligned} \tag{3-15}$$

Como A é normal, a eq. 3-15 mostra que, para cada $y \in H$, $NN^*y = N^*Ny$.

Então, N é normal e, conseqüentemente,

$$\begin{aligned}\langle Ny; Ny \rangle &= \langle y; N^*Ny \rangle \\ &= \langle y; NN^*y \rangle \\ &= \langle N^*y; N^*y \rangle.\end{aligned}\tag{3-16}$$

Por conseguinte,

$$\|A^*(A - \lambda I)y\| = \|(A^* - \bar{\lambda}I)Ay\|.\tag{3-17}$$

Como (λ, x) é um autopar de A , $(A - \lambda I)x = 0$. Logo, a eq. 3-17 leva a $(A^* - \bar{\lambda}I)Ax = 0$, ou, equivalentemente, $AA^*x = A^*Ax = \bar{\lambda}Ax = |\lambda|^2x$, isto é, $(|\lambda|^2, x)$ é um autopar de AA^* . ■

Portanto, o operador AA^* é auto-adjunto, não-negativo e, além disso, toda autofunção de A também é uma autofunção de AA^* . Assim, é possível adotar uma metodologia semelhante àquela apresentada para operadores auto-adjuntos indefinidos: resolve-se $AA^*x = \theta x$ e, depois, verifica-se se x é autofunção de A utilizando o teste sugerido na eq. 3-11. Finalmente, se x for uma autofunção de A , deve-se empregar a *razão de Rayleigh* para calcular o autovalor associado à x .

Com isso, a metodologia de solução de equações de autovalores usando PG está definida para qualquer operador normal. Na próxima seção, será abordado o caso de operadores que não são normais.

3.2.5

Propostas de Adaptações para Outros Operadores

Embora os operadores normais englobem uma grande quantidade de problemas práticos, eles não são úteis em diversos problemas de autovalores. Por exemplo, este trabalho mostrará em um capítulo subsequente o estudo de caso referente a poços de petróleo, no qual deseja-se encontrar modos de propagação em um guia de ondas radialmente estratificado. Neste caso, é possível mostrar que o operador diferencial não é normal.

Encontrar uma técnica capaz de achar um número arbitrário de autopares de um operador diferencial qualquer é uma tarefa enorme e provavelmente ineficaz. Isso ocorre pelas razões abaixo descritas.

1. *Não necessariamente o conjunto de autopares de um operador diferencial é discreto.* Por exemplo, o operador derivada definido para qualquer função uma vez diferenciável tem como espectro o conjunto não enumerável

de autopares $\{(u; f(x)) : u \in \mathbb{C}, f(x) = e^{ux}, x \in \mathbb{C}\}$. Isto ocorre porque $e^{ux}_{[x]} = ue^{ux}$ para todo $u \in \mathbb{C}$.

2. Um método genérico não exploraria características particulares do operador em questão. Isso provavelmente tornaria o método ineficiente, ou seja, para cada classe particular de operadores, seria possível encontrar um método que explorasse características particulares desta classe e, assim, seria mais eficiente, mais robusto e provavelmente mais simples de implementar. Por exemplo, o modelo apresentado anteriormente utiliza diversas propriedades de operadores auto-adjuntos, como a presença de espectro discreto, autovalores reais e o fato do mínimo da *razão de Rayleigh* ser igual ao menor autovalor.

Portanto, não é pretensão deste trabalho desenvolver uma metodologia rigorosa de determinação de um número arbitrário de autopares de um operador diferencial qualquer. Sendo assim, esta seção mostra uma heurística para abordar o problema de autovalores para o caso mais geral.

No caso geral de um operador linear, a *razão de Rayleigh* continua sendo útil para calcular autovalores a partir de autovetores. Contudo, como os autovalores são, em geral, complexos, não faz mais sentido minimizar a razão de Rayleigh. Neste caso, o problema de minimização pode ser escrito de duas formas:

1. minimiza-se $\|Af - F_A[f]f\| / (|F_A[f]| \cdot \|f\|)$ sujeito às restrições que f deve satisfazer. Se f for uma autofunção de A , então $F_A[f]$ é o autovalor associado à f e, portanto, $\|Af - F_A[f]f\| = 0$, que é o valor mínimo da norma. Por outro lado, se $\|Af - F_A[f]f\| = 0$, então $Af = F_A[f]f$ e, por definição, $(F_A[f]; f)$ é um autopar de A . Assim, $\|Af - F_A[f]f\| = 0$ se, e somente se, $(F_A[f]; f)$ for uma solução para o problema de autovalores. Note, contudo, que não há garantia que $|F_A[f]|$ seja o menor possível.
2. utiliza-se um conjunto finito de Q funções de teste t_j e calcula-se o conjunto de erros $\epsilon_j = (|F_A[f]| \cdot \|f\|)^{-1} \cdot \langle Af - F_A[f]f; t_j \rangle$ para $j = 1, 2, \dots, Q$. Posteriormente, minimiza-se a soma dos quadrados dos erros $\sum_j \bar{\epsilon}_j \epsilon_j$ sujeito às restrições que f deve satisfazer. Idealmente, o conjunto de funções de teste deve ser ortogonal e estar contido em $L^2(\Omega)^N$, assim como f . Exemplos de funções de teste com essa propriedade são as funções base das séries de Fourier, como senos, cossenos e polinômios ortogonais.

Finalmente, é importante notar que a metodologia para encontrar outros autovalores baseada na ortogonalização de Gram-Schmidt, apresentada anteri-

ormente, não é geralmente aplicável no caso de qualquer operador linear. Isto ocorre porque no caso geral de um operador linear não há garantias de que exista um conjunto ortogonal de autovetores. Por esta razão, as próximas linhas estabelecem uma nova metodologia para encontrar os demais autovetores, dado um conjunto linearmente independente de autovetores conhecidos.

No caso mais geral em que o operador não tem autovetores ortogonais, deseja-se apenas que os novos autovetores encontrados sejam linearmente independentes aos autovetores já conhecidos. Para isso, proponho definir um operador de projeção, que “remove” as componentes dos vetores conhecidos do conjunto de busca, garantindo que todos os vetores presentes no espaço de busca sejam linearmente independentes em relação aos autovetores conhecidos. Este operador é descrito nos próximos parágrafos.

Definição 3.6 (Projeção de remoção de autovetores) *Seja H um espaço de Hilbert com produto interno $\langle \cdot; \cdot \rangle$ e A um operador linear com domínio e contradomínio iguais a H . Considere que $S = \{v_1, v_2, \dots, v_N\} \subseteq H$ é um conjunto de autovetores linearmente independentes de A . Nesse contexto, este trabalho define a projeção P para cada vetor $u \in H$ como*

$$Pu = u - \sum_{i=1}^N a_i v_i, \quad (3-18)$$

onde os coeficientes a_i são soluções do sistema linear

$$\begin{bmatrix} \langle v_1; v_1 \rangle & \langle v_2; v_1 \rangle & \dots & \langle v_N; v_1 \rangle \\ \langle v_1; v_2 \rangle & \langle v_2; v_2 \rangle & \dots & \langle v_N; v_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_1; v_N \rangle & \langle v_2; v_N \rangle & \dots & \langle v_N; v_N \rangle \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} \langle v_1; u \rangle \\ \langle v_2; u \rangle \\ \vdots \\ \langle v_N; u \rangle \end{bmatrix} \quad (3-19)$$

Teorema 3.7 (Propriedades da projeção) *Considere a projeção definida em 3.6 e seja $z \in H$ tal que $z = Pu$, $u \in H$. Então, as seguintes afirmações são verdadeiras:*

1. o conjunto $\{z\} \cup S$ é linearmente independente.
2. z é o vetor linearmente independente dos vetores de S que tem a menor norma.

Prova. Para provar a afirmação 1, observe que a equação

$$z + \sum_{i=1}^N b_i v_i = 0 \quad b_i \in \mathbb{C}, i = 1, 2, \dots, N. \quad (3-20)$$

pode ser reescrita como

$$\sum_{i=1}^N (b_i - a_i) \langle v_i; v_j \rangle = - \langle u; v_j \rangle \quad j = 1, 2, \dots, N. \quad (3-21)$$

Contudo, por definição

$$\sum_{i=1}^N a_i \langle v_i; v_j \rangle = \langle u; v_j \rangle \quad j = 1, 2, \dots, N. \quad (3-22)$$

Consequentemente, a eq. 3-21 é equivalente a

$$\sum_{i=1}^N b_i \langle v_i; v_j \rangle = 0 \quad j = 1, 2, \dots, N, \quad (3-23)$$

cuja única solução é $b_i = 0$ para todo $i = 1, 2, \dots, N$, pois o conjunto S é linearmente independente. Assim, a eq. 3-20 não apresenta soluções não-triviais, isto é, z não pode ser escrito como combinação linear dos elementos de S , o que prova a afirmação 1.

Para provar a afirmação 2, seja $Y = \|z\|^2$. Pela definição da projeção, $Y = \left\| u - \sum_{i=1}^N a_i v_i \right\|^2$. Logo, utilizando as propriedades de produto interno, nota-se que

$$\begin{aligned} Y &= \left\| u - \sum_{i=1}^N a_i v_i \right\|^2 \\ &= \left\langle u - \sum_{i=1}^N a_i v_i; u - \sum_{j=1}^N a_j v_j \right\rangle \\ &= \left\langle u; u - \sum_{j=1}^N a_j v_j \right\rangle - \sum_{i=1}^N a_i \left\langle v_i; u - \sum_{j=1}^N a_j v_j \right\rangle \\ &= \langle u; u \rangle - \sum_{j=1}^N \bar{a}_j \langle u; v_j \rangle - \sum_{i=1}^N a_i \langle v_i; u \rangle + \sum_{i=1}^N \sum_{j=1}^N a_i \bar{a}_j \langle v_i; v_j \rangle \\ &= \langle u; u \rangle - 2 \sum_{i=1}^N \operatorname{Re} (a_i \langle v_i; u \rangle) + \sum_{i=1}^N \sum_{j=1}^N a_i \bar{a}_j \langle v_i; v_j \rangle. \end{aligned} \quad (3-24)$$

Portanto,

$$\begin{aligned} \frac{dY}{da_k} &= -2 \operatorname{Re} \langle v_k; u \rangle + \sum_{j=1}^N \bar{a}_j \langle v_k; v_j \rangle + \sum_{i=1}^N a_i \langle v_i; v_k \rangle \\ &= -2 \operatorname{Re} \langle v_k; u \rangle + 2 \sum_{i=1}^N \operatorname{Re} (a_i \langle v_i; v_k \rangle). \end{aligned} \quad (3-25)$$

Contudo, pela definição 3.6, $\sum_{i=1}^N \operatorname{Re} (a_i \langle v_i; v_k \rangle) = \operatorname{Re} \langle v_k; u \rangle$ para cada $k =$

$1, \dots, N$. Sendo assim, $dY/da_k = 0$ para todo $k = 1, \dots, N$. Além disso,

$$\frac{d^2Y}{da_m da_k} = 2\text{Re} \langle v_m; v_k \rangle \quad k, m = 1, 2, \dots, N \quad (3-26)$$

ou seja, a forma quadrática $a^H \Pi a$, onde Π é a Hessiana e $a^H = [\bar{a}_1, \dots, \bar{a}_N]$, é dada por:

$$\begin{aligned} a^H \Pi a &= \sum_{m=1}^N \sum_{k=1}^N a_m 2\text{Re} \langle v_m; v_k \rangle \bar{a}_k \\ &= 2\text{Re} \left\langle \sum_{m=1}^N a_m v_m; \sum_{k=1}^N a_k v_k \right\rangle \\ &= 2\text{Re} \left\| \sum_{m=1}^N a_m v_m \right\|^2 \\ &\geq 0. \end{aligned} \quad (3-27)$$

Portanto, a matriz Hessiana é positiva semidefinida e, com isso, o ponto (a_1, a_2, \dots, a_N) é ponto de mínimo, o que prova a afirmação 2. ■

Portanto, a aplicação de P a qualquer vetor $u \in H$ gera um vetor que é linearmente independente dos autovetores já conhecidos. Além disso, este vetor é o vetor mais próximo ao conjunto formado por todas as combinações lineares dos vetores de S ($\text{span}(S)$). Conseqüentemente, no caso particular de $u \in \text{span}(S)$, $Pu = 0$, ou seja, a projeção de um vetor linearmente *dependente* aos vetores de S é o vetor nulo. Assim, intuitivamente qualquer “componente” de u que tenha alguma dependência com autovetores conhecidos é removida em Pu .

Com a projeção 3.6 em mãos, basta aplicar a mesma ideia apresentada para operadores não-negativos: dado o conjunto S de autovetores conhecidos, constrói-se a projeção 3.6 e resolve-se o problema de autovalores

$$APx = \lambda Px. \quad (3-28)$$

O processo é repetido quantas vezes for necessário, até que se encontre uma quantidade suficientemente grande de autopares de A .

Com isso, o modelo de determinação de autopares está completamente descrito para todos os tipos de operadores diferenciais lineares. Na próxima seção, será mostrado como utilizar o modelo proposto para a solução de problemas de autovalores paramétricos.

3.2.6

Modelo de Solução Paramétrico

Resta, então, explicar como tornar paramétrico o modelo descrito nas últimas seções. Portanto, esta seção explica como resolver o problema de autovalores parametrizado $A_p f_p = \lambda_p f_p$, onde $p \in C \subseteq \mathbb{C}^L$ é um parâmetro pré-especificado L -dimensional.

Neste tipo de problema, o operador diferencial tipicamente tem sua expressão matemática dependente de p . O capítulo 2 deu um exemplo utilizando o operador de Laplace. Como o operador depende de p , é de se esperar que seus autopares também sejam dependentes de p . Sendo assim, o modelo proposto agora deve considerar a presença do parâmetro p e fornecer como resposta um conjunto de autofunções cuja expressão analítica é parametrizada por p . Dessa forma, utilizando a razão de Rayleigh, determina-se o autovalor a partir de f_p , para cada valor do parâmetro p .

Um ponto importante (e bem interessante) é que geralmente o valor do parâmetro p impacta na definição do domínio Ω_p da função f_p e, com isso, o produto interno passa a depender de p . O mesmo ocorre com a norma derivada do produto interno: $\|v\|_p = \sqrt{\langle v; v \rangle_p}$ para todo $v \in L^2(\Omega_p)^N$.

Feitas as considerações, nesta tese mostro como utilizar o modelo de solução já proposto para encontrar os autovalores da família de operadores parametrizada A_p , assumindo que A_p é um operador auto-adjunto não negativo para todo p . Caso A_p não satisfaça esta restrição, as modificações apresentadas anteriormente podem ser aplicadas para converter o operador em um operador não-negativo, conservando o espectro.

O modelo proposto pode ser adaptado para encontrar autofunções cujo autovalor associado seja o menor do espectro de A_p , para todo $p \in C \subseteq \mathbb{C}^L$. Para isso, proponho discretizar o conjunto C com uma malha finita $\tilde{C} \in C$ e resolver o problema de otimização

$$\begin{aligned} & \min \sum_{p \in \tilde{C}} F_{A_p}[f_p] \\ & \text{sujeito a} \\ & B_p f_p = 0 \text{ para todo } p \in \tilde{C} \\ & f_p \in L^2(\Omega_p)^N \text{ para todo } p \in \tilde{C}. \end{aligned} \quad (3-29)$$

Neste problema, o funcional de Rayleigh parametrizado por $p \in C$ é calculado

para cada valor de $p \in C$ como

$$F_{A_p}[v] = \frac{\langle A_p v; v \rangle_p}{\langle v; v \rangle_p}$$

$$v \in L^2(\Omega_p)^N \tag{3-30}$$

Além disso, o somatório é tomado em uma malha finita de pontos contida em C .

Com isso, aplica-se a metodologia já explicada para operadores não-negativos, modificando a função objetivo para considerar a nova formulação. O resultado final é uma autofunção que contém o parâmetro p como argumento e está associada ao menor autovalor, dado p . Para encontrar os demais autovalores, a metodologia explicada para o caso não-paramétrico continua valendo, com o detalhe que as projeções dependem do parâmetro p . Mais uma vez, se o operador A_p não for não-negativo, deve-se modifica-lo para torna-lo não negativo, conforme explicado nas seções sobre operadores normais.

Finalmente, caso o operador A_p não seja normal, o problema de otimização deve ser escrito como

$$\min \sum_{p \in \tilde{C}} \frac{\|A_p f - F_{A_p}[f]f\|}{\|F_{A_p}[f_p]f_p\|}$$

sujeito a

$$B_p f_p = 0 \text{ para todo } p \in \tilde{C}$$

$$f_p \in L^2(\Omega_p)^N \text{ para todo } p \in \tilde{C}. \tag{3-31}$$

Nos próximos capítulos, serão vistos exemplos de aplicação desta técnica, com resultados comparativos com outras técnicas, como o Método dos Elementos Finitos.

4

Estudos de Caso: Operadores Normais

Este capítulo foca em responder à pergunta “quais são os resultados?”. Para isso, este capítulo apresenta os resultados obtidos com o modelo descrito no capítulo 3 para problemas de autovalores de operadores diferenciais normais.

As próximas seções apresentam estudos de caso clássicos de Física e Engenharia. Inicialmente, formula-se o problema como um problema de autovalores de um operador diferencial e, posteriormente, demonstra-se que este operador é normal. Na sequência, apresenta-se a solução analítica, as soluções por Método das Diferenças Finitas e Método dos Elementos Finitos (quando aplicável) e a solução obtida a partir do modelo proposto neste trabalho, comparando-as em termos de acurácia e tempo de execução. Finalmente, analisa-se a eficiência e o que pode ser melhorado no modelo proposto para cada estudo de caso.

4.1

Estudo de Caso: Operador de Laplace

4.1.1

Definição

O operador escalar de Laplace $A : L^2(\Omega) \rightarrow L^2(\Omega)$ é definido para funções quadrado-integráveis com domínio $\Omega \subseteq \mathbb{R}^2$ que são nulas no contorno. Formalmente,

$$\begin{aligned} A &= \nabla^2 f = f_{[xx]} + f_{[yy]} \\ D(A) &= \{f \in L^2(\Omega) : \nabla^2 f \in L^2(\Omega), \\ & f(t) = 0 \text{ se } t \in \Gamma\} \\ \Gamma &\text{ é o contorno de } \Omega \\ \Omega &\subset \mathbb{R}^2. \end{aligned} \tag{4-1}$$

Este operador é auto-adjunto [4, p. 160] sob o produto interno

$$\begin{aligned} \langle f; g \rangle &= \int_{\Omega} \bar{g} f d\Omega \\ f, g &\in L^2(\Omega), \end{aligned} \tag{4-2}$$

onde a barra superior denota o complexo conjugado. A demonstração envolve o segundo teorema de Green e o fato de as funções do domínio do operador serem nulas no contorno.

Além disso, o operador de Laplace conforme definido em 4-1 é não-positivo, fato que também pode ser demonstrado pelo segundo teorema de Green:

$$\begin{aligned}
 \langle Af; f \rangle &= \int_{\Omega} \bar{f} \nabla^2 f \, d\Omega \\
 &= \int_{\Gamma} \bar{f} f_{[n]} \, d\Gamma - \int_{\Omega} \overline{\nabla f} \cdot \nabla f \, d\Omega \\
 &= - \int_{\Omega} |\nabla f|^2 \, d\Omega \\
 &\leq 0, \\
 &\text{para todo } f \in D(A). \tag{4-3}
 \end{aligned}$$

Portanto, o operador $-A$ com A definido por 4-1 é auto-adjunto e não-negativo, se encaixando então nos pré-requisitos do modelo proposto.

4.1.2

Experimento: domínio retangular com solução analítica

Nesta seção, estamos interessados em resolver o problema de autovalores

$$\begin{aligned}
 -Af &= \lambda f \\
 (Bf)(x, y) &= \begin{cases} f(x, y) & (x, y) \in \Gamma \\ 0 & \text{caso contrário} \end{cases} \\
 Bf &= 0 \\
 \Omega &= [0; 1]^2 \\
 \Gamma &\text{ é o contorno de } \Omega. \\
 \lambda &> 0 \\
 f &\in D(A). \tag{4-4}
 \end{aligned}$$

Ou seja, o domínio das funções é o quadrado unitário no \mathbb{R}^2 . Note que o operador B é construído de modo a forçar que a condição $Bf = 0$ implique em $f = 0$ no contorno Γ . Este tipo de condição de contorno é chamado de condição de contorno de Dirichlet.

Este tipo de problema de autovalores é muito comum na Física e na Engenharia. Um exemplo é a determinação de modos transversos elétricos em guias de ondas retangulares de paredes metálicas. Nestes guias, as constantes de onda dos modos de propagação estão associadas aos autovalores e as ondas

eletromagnéticas referentes a cada modo estão relacionadas às autofunções.

Este problema foi escolhido como caso de estudo inicial por ter solução analítica. Em particular, para um quadrado de lado a , os autopares são dados por [38]:

$$\begin{aligned} \lambda_{n,m} &= (m^2 + n^2)\pi^2 \\ f_{n,m}(x, y) &= \text{sen}\left(\frac{n\pi x}{a}\right) \text{sen}\left(\frac{m\pi y}{a}\right) \\ n, m &= 1, 2, 3, 4, \dots \\ (x; y) &\in [0; a]^2 \\ a &> 0. \end{aligned} \tag{4-5}$$

O menor autovalor é $\lambda = 2\pi^2$, cuja autofunção é $f_{1,1}$. Além disso, temos $\lambda_{1,1} \leq \lambda_{2,0} \leq \lambda_{0,2} \leq \dots$. Note também que, se f é uma autofunção, então bf onde $b \in \mathbb{C} - \{0\}$ e é uma constante também é uma autofunção associada ao mesmo autovalor de f .

Para resolver este problema de autovalores utilizando o modelo proposto, deve-se primeiramente definir o número de árvores usadas para representar o indivíduo como na equação 3-3. Neste caso, motivado pela método da separação de variáveis para solução de equações diferenciais parciais, proponho escrever $f(x, y) = f_X(x)f_Y(y)$ e, portanto, utilizar 2 conjuntos de k_a árvores sem superposição para representar cada uma das funções f_X e f_Y . Então, considerando 3-3, $f_X = \sum_{l=1}^{k_a} d_{X,l}f_{X,l}$ e $f_Y = \sum_{l=1}^{k_a} d_{Y,l}f_{Y,l}$. Para determinar os coeficientes, utiliza-se a metodologia explicada no algoritmo 1, isto é, escolhe-se o ponto $(d_{X,1}, \dots, d_{X,k_a}, d_{Y,1}, \dots, d_{Y,k_a})$ que minimiza a *razão de Rayleigh* $F_A[f]$ sujeito a $f(x, y) = 0$ se $(x; y) \in \Gamma$. Mais especificamente, queremos resolver o problema de otimização

$$\begin{aligned} &\mathbf{min} F_A[f] \\ &\text{sujeito a} \\ &f(0, y) = 0 \rightarrow \sum_{l=1}^{k_a} d_{X,l}f_{X,l}(0) = 0 \\ &f(1, y) = 0 \rightarrow \sum_{l=1}^{k_a} d_{X,l}f_{X,l}(1) = 0 \\ &f(x, 0) = 0 \rightarrow \sum_{l=1}^{k_a} d_{Y,l}f_{Y,l}(0) = 0 \\ &f(x, 1) = 0 \rightarrow \sum_{l=1}^{k_a} d_{Y,l}f_{Y,l}(1) = 0. \end{aligned} \tag{4-6}$$

Note que o problema 4-6 apresenta restrições lineares nas variáveis $(d_{X,1}, \dots, d_{X,k_a}, d_{Y,1}, \dots, d_{Y,k_a})$. Como o problema 4-6 é resolvido a cada avaliação de cada indivíduo em particular, as funções $f_{X,l}$ e $f_{Y,l}$ são conhecidas, pois fazem parte do cromossomo do modelo de programação genética. Portanto, para resolver 4-6, basta utilizar algum algoritmo eficiente de otimização de funções reais com restrições lineares, como o BFGS Quasi Newton [34, 35].

Os parâmetros utilizados no modelo de programação de genética estão especificados na tabela 1. Primeiramente, resolveu-se o problema não-paramétrico com $a = 1$. Para garantir que as restrições em 4-6 sempre tenham solução, temos que ter $k_a \geq 3$.

Tabela 1: Modelo proposto para achar os autopares do operador de Laplace.

População	30 indivíduos
Gerações	500 gerações
Núm. Árvores (k_a)	de 3 a 6 (vários experimentos)
Taxa de mutação	fixa em 10%
Tipos de mutação	6 tipos descritos no cap. 3
Tipo de recombinação	sub-árvore
Funções	$\{\text{sen}(t), \text{cos}(t), *, +\}$
Terminais	$\{t, \pi, 0, 1, -1\}$
Otimizador Local	<i>BFGS Quasi-Newton</i> [34, 35]
Avaliação	alg. 1 + derivadas e simplificação simbólicas
Profundidade Máxima	4 níveis por árvore

Antes de apresentar os resultados obtidos, é preciso enfatizar que o modelo proposto é não-determinístico, pois a geração da população inicial é aleatória, assim como a seleção de indivíduos para reprodução e a aplicação de operadores. Por isso, os resultados podem ser diferentes ao rodar o modelo várias vezes para resolver o mesmo problema. Contudo, espera-se que, embora diferentes, os resultados sejam próximos ou equivalentes. Além disso, espera-se também que o problema seja resolvido na maioria das rodadas. Para testar este fato, foram rodados 40 experimentos independentes, isto é, o modelo foi aplicado 40 vezes ao mesmo problema para analisar a resposta.

Em um primeiro momento, deseja-se encontrar o menor autovalor e a autofunção associada. Esta primeira etapa executa o algoritmo 2 por 40 vezes (são 40 rodadas) utilizando a função de avaliação definida no algoritmo 1 para encontrar o menor autovalor e autofunção em 40 oportunidades. Este experimento foi repetido variando o valor de k_a de 3 a 6. A tabela 2 mostra quantas vezes foi encontrada a autofunção $f_{1,1}$ ou qualquer função múltipla dela. Além disso, a tabela também mostra quantas vezes nenhuma função foi encontrada.

Tabela 2: Taxa de sucesso do modelo proposto para achar o menor autopar do operador de Laplace.

Num. Árvores (k_a)	$f_{1,1}$	$bf_{1,1}$	Outra
3	27	10	3
4	21	12	7
5	22	14	4
6	19	10	11

Pelos resultados da tabela 2, é visível que o aumento do número de árvores do indivíduo aumenta, também, a taxa de rodadas na qual a solução exata não é encontrada (coluna “Outra”). Isso ocorre porque o aumento do cromossomo implica em aumento do espaço de busca, tornando o problema mais difícil de resolver. Além disso, percebe-se que, fixado o número de árvores, a chance de encontrar a autofunção mais simples é maior do que a chance de encontrar a autofunção mais complexa. Esse é um comportamento esperado que ocorre porque o modelo proposto força uma pressão de seleção maior para indivíduos mais simples (com menos nós).

Na figura 5, apresento a evolução média do melhor indivíduo nos 40 experimentos. É possível observar que aumentar o número de árvores torna a evolução mais lenta, i.e., leva mais gerações para que o melhor indivíduo atinja o valor mínimo de $2\pi^2$. Esse é o comportamento esperado porque o aumento do número de árvores implica em aumento do espaço de busca, o que dificulta o problema de otimização. Uma forma de contornar esta característica

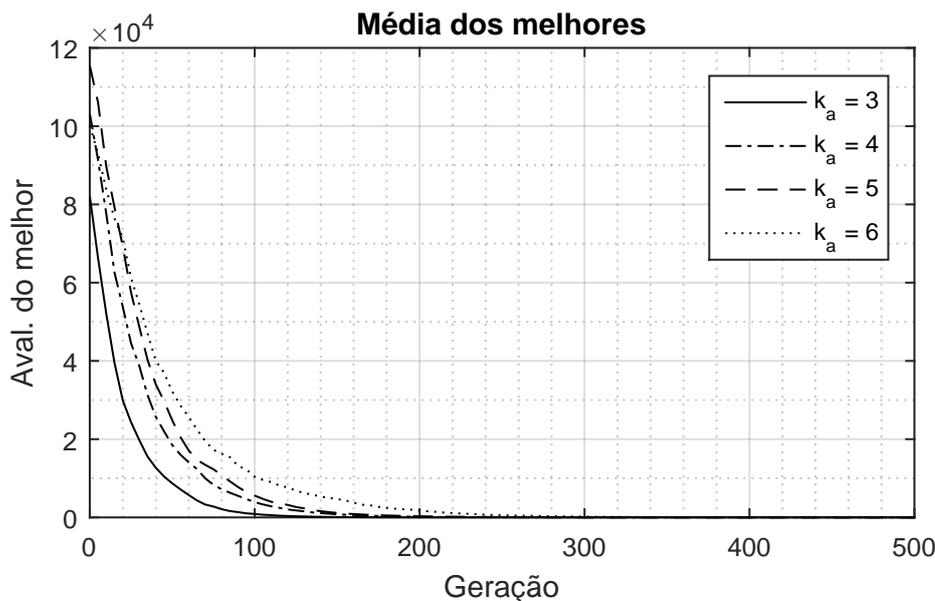


Figura 5: Média do melhor indivíduo dentre os 40 experimentos para cada quantidade de árvores para encontrar o menor autovalor.

é aumentar o número total de gerações, a custo de aumentar o tempo gasto na otimização.

A partir da tabela 2 e da figura 5, conclui-se que o modelo funciona com robustez para encontrar o menor autovalor do operador de Laplace. O próximo passo é utilizar a autofunção encontrada para encontrar o segundo menor autovalor e assim sucessivamente.

A título de experimento, foi assumido que $f_{1,1}$ é conhecida e aplicou-se a modificação sugerida no capítulo 3 para buscar o segundo autovalor e sua autofunção não nula, que no caso são, respectivamente, $5\pi^2$ e $f_{2,1}$ ou $f_{1,2}$. Novamente, a tabela 3 mostra os resultados após 40 rodadas.

Tabela 3: Taxa de sucesso do modelo proposto para achar o 2o. autpar do operador de Laplace.

Num. Árvores	$f_{2,1}$ ou $f_{1,2}$	$b_1 f_{2,1} + b_2 f_{1,2}$	Outra
3	19	16	5
4	13	15	12
5	9	11	20
6	5	12	23

Novamente, verifica-se que o método encontrou a solução exata em todos os casos, com uma taxa maior para menores quantidades de árvores. Percebeu-se também que a pressão de seleção por indivíduos com menos nós de fato funcionou como esperado, porque em geral encontrou-se as soluções mais simples.

É preciso enfatizar que, na prática, o usuário do método *não precisa rodar a técnica por 40 vezes* para obter um resultado confiável. Na realidade, esta análise é útil apenas para entender a taxa de sucesso do algoritmo, dado o número de gerações (neste caso, 500). Na prática, todos os algoritmos de otimização não-linear dependem de inicializações e, por isso, podem demandar mais ou menos iterações para convergir para a solução ótima. No caso do modelo proposto, as gerações são o equivalente às iterações e a população inicial equivale ao ponto inicial fornecido aos algoritmos de otimização não-linear. Portanto, acredita-se que o modelo proposto *sempre encontrará a solução, desde que um número suficientemente grande de gerações seja utilizado*.

Para ilustrar esta argumentação, podemos analisar novamente a figura 5 e observar que, na prática, a avaliação tem uma tendência decrescente em todos os experimentos feitos. Isso mostra que, mesmo que o valor mínimo não seja encontrado nas 500 gerações, ele irá ser encontrado com um número suficientemente grande de gerações.

Portanto, para garantir que a metodologia proposta sempre encontre a solução, basta modificar o critério de parada de modo a oferecer garantias

melhores. Assim, no caso do problema de autovalores desta seção, se definirmos o erro $\epsilon = \|-Af - \lambda f\| \cdot \|f\|^{-1}$, podemos evoluir até que $\epsilon \leq \epsilon_m$, onde ϵ_m é uma constante pequena, como 10^{-5} . Portanto, ao invés de parar a heurística de busca por número de gerações, estaríamos parando quando esta encontrasse uma solução com erro aceitavelmente pequeno. Naturalmente, o número de gerações resultante dependerá de quão boa é a população inicial, assim como os algoritmos de otimização não-linear tradicionais dependem da inicialização fornecida pelo usuário.

Para concluir esta seção, vamos analisar os resultados para solução do problema paramétrico, isto é, queremos aplicar o modelo proposto para resolver o problema paramétrico

$$\begin{aligned}
 -Af &= \lambda f \\
 (B_a f)(x, y) &= \begin{cases} f(x, y) & (x, y) \in \Gamma_a \\ 0 & \text{caso contrário} \end{cases} \\
 &= 0 \\
 \Omega_a &= [0; a]^2 \\
 \Gamma_a &\text{ é o contorno de } \Omega_a. \\
 \lambda &> 0 \\
 f &\in L^2(\Omega_a) \\
 a &\in [0.5; 1.5]
 \end{aligned} \tag{4-7}$$

Neste problema, a solução é função do parâmetro a e é dada pela equação 4-5. O objetivo agora é adotar a metodologia de solução paramétrica apresentada no capítulo 3 para encontrar as equações 4-5.

Para resolver este problema paramétrico, utilizou-se uma malha uniforme de 1000 pontos no intervalo $[0.5; 1.5]$, incluindo seus extremos. Com isso, o modelo de solução paramétrico apresentado no capítulo 3 foi usado para resolver o problema

$$\begin{aligned}
 \min_{a \in [0.5; 1.5]} & \sum F_{\nabla^2}[f_a] \\
 \text{sujeito a} & \\
 B_a f_a &= 0 \text{ para todo } a \in [0.5; 1.5] \\
 f_a &\in L^2(\Omega_a) \text{ para todo } a \in [0.5; 1.5],
 \end{aligned} \tag{4-8}$$

onde

$$\begin{aligned} \langle f; g \rangle_a &= \int_0^a \int_0^a \overline{g(x, y)} f(x, y) \, dx dy \\ \|f\|_a &= \sqrt{\langle f; f \rangle_a} \\ f, g &\in L^2(\Omega_a) \end{aligned} \tag{4-9}$$

e $F_{\nabla^2}[f_a]$ é o funcional de Rayleigh referente ao operador Laplaciano, também parametrizado pelo valor de a :

$$\begin{aligned} F_{\nabla^2}[g] &= \frac{\langle \nabla^2 g; g \rangle_a}{\langle g; g \rangle_a} \\ g &\in L^2(\Omega_a). \end{aligned} \tag{4-10}$$

A solução deste problema entrega para cada ponto a na malha de pontos uma autofunção parametrizada pelo valor de a e cujo autovalor é o menor possível. Uma possível forma de implementar 4-8 é representar a função f_a utilizando 4 funções univariadas: $f_a(x, y) = f_X(x)f_Y(y)g_X(a - x)g_Y(a - y)$. Cada uma das 4 funções utiliza k_a árvores e é nula quando seu argumento é nulo. Dessa forma, $f_a(a, y) = f_a(0, y) = f_a(x, a) = f_a(x, 0) = 0$ para cada $a \in [0.5; 1.5]$. Portanto, o problema 4-8 passa a ser escrito como

$$\begin{aligned} \min \quad & \sum_{a \in [0.5; 1.5]} F_{\nabla^2}[f_a] \\ \text{sujeito a} \quad & \\ f(0, y) = 0 \rightarrow & \sum_{l=1}^{k_a} d_{X,l} f_{X,l}(0) = 0 \\ f(a, y) = 0 \rightarrow & \sum_{l=1}^{k_a} b_{X,l} g_{X,l}(0) = 0 \\ f(x, 0) = 0 \rightarrow & \sum_{l=1}^{k_a} d_{Y,l} f_{Y,l}(0) = 0 \\ f(x, a) = 0 \rightarrow & \sum_{l=1}^{k_a} b_{Y,l} g_{Y,l}(0) = 0. \end{aligned} \tag{4-11}$$

Neste caso, a letra b foi utilizada para designar coeficientes das funções g_X e g_Y .

Para analisar a qualidade da solução de uma determinada rodada da heurística, define-se a pontuação da solução como sendo o número de pontos da malha de pontos para o parâmetro a cuja solução é exata. Considera-se que uma solução normalizada f_{sol} é exata se $\|f_{sol} - f_{exa}\|_a \leq 10^{-8}$, onde f_{exa} é a solução exata analítica que se espera encontrar, após normalização

($\|f_{exa}\|_a = 1$ para algum $a \in [0.5; 1.5]$). Note que f_{sol} também é normalizada, ou seja, $\|f_{sol}\|_a = 1$ para algum $a \in [0.5; 1.5]$.

A tabela 4 mostra a quantidade de rodadas que superaram determinada pontuação.

Tabela 4: Quantidade de rodadas cuja pontuação da solução é superior a determinado percentual para encontrar $f_{1,1}$ (colunas).

Num. Árvores	70%	80%	90%
2	40	37	33
3	37	34	28
4	34	30	24
5	29	21	19

Nota-se que a solução é exata para mais de 90% dos pontos em 31 das 40 rodadas executadas com o cromossomo de 2 árvores. Isso mostra que o problema paramétrico é mais difícil, mas o modelo proposto consegue resolvê-lo. Além disso, vale lembrar que aumentar o número de gerações tende a melhorar a taxa de acerto do modelo proposto, como ocorre em qualquer algoritmo de otimização.

Nos experimentos apresentados nesta seção, os modelos com mais árvores apresentaram degradação devido ao aumento do espaço de busca e, principalmente, devido ao fato de a solução exata *não ter somatórios ou combinações lineares*. No entanto, espera-se que o modelo multi-árvore tenha desempenho melhor em problemas cujas autofunções sejam representadas por somas de funções mais elementares, justificando, assim, o aumento do espaço de busca.

4.1.3

Experimento: Domínio senoidal

O operador de Laplace pode ser estudado para diversos formatos de domínio. Por exemplo, em [39] o autor estuda os autopares do operador de Laplace assumindo domínio triangular e expressa as autofunções analiticamente. Contudo, não é interessante neste trabalho adotar outro estudo de caso com solução analítica. Nesta seção, a presente tese foca em comparar a técnica proposta com o Método dos Elementos Finitos e o Método das Diferenças Finitas, utilizando um estudo de caso projetado para estressar ao máximo as técnicas de solução, de modo a compara-las em termos de acurácia e tempo de execução.

O problema abordado nesta seção foi construído a partir do problema apresentado na seção anterior, considerando que uma das arestas do domínio retangular fora substituída por uma senóide de frequência k e amplitude α , conforme mostra a figura 6.

Nos limites do pseudo-quadrado S , as funções são nulas. Assim, $f(0, y) = f(1, y) = 0$, para todo $y \in [0; 1]$ e $f(x, 1) = f(x, \alpha \text{sen}(2\pi kx)) = 0$, para todo $x \in [0; 1]$. Portanto, o problema de autovalores (não-paramétrico) é

$$\begin{aligned}
 -Af &= \lambda f \\
 (Bf)(x, y) &= \begin{cases} f(x, y) & (x, y) \in \Gamma \\ 0 & \text{caso contrário} \end{cases} \\
 &= 0 \\
 S &= [0; 1]^2 - \{(x; y) : 0 \leq x \leq 1, y \leq \alpha \text{sen}(2\pi kx)\} \\
 \Gamma &\text{ é o contorno de } S. \\
 \lambda &> 0 \\
 f &\in L^2(S), \tag{4-12}
 \end{aligned}$$

enquanto que o problema paramétrico (a é o parâmetro) é dado para uma

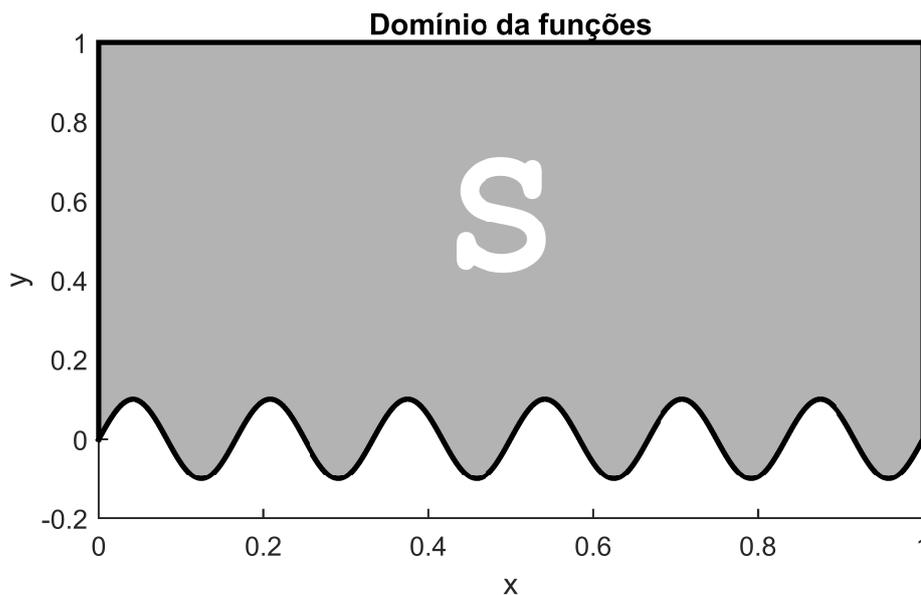


Figura 6: Domínio S das funções do operador de Laplace para o estudo de caso da seção 4.1.3. A senóide tem frequência $k = 6$.

frequência k por

$$\begin{aligned}
 -Af_a &= \lambda f_a \\
 (B_a f)(x, y) &= \begin{cases} f_a(x, y) & (x, y) \in \Gamma_a \\ 0 & \text{caso contrário} \end{cases} \\
 &= 0 \\
 S_a &= [0; a]^2 - \left\{ (x; y) : 0 \leq x \leq a, y \leq \alpha \operatorname{sen}\left(\frac{2\pi kx}{a}\right) \right\} \\
 \Gamma_a &\text{ é o contorno de } S_a. \\
 \lambda &> 0 \\
 k &\in \{1, 2, 3, \dots\} \\
 f &\in L^2(S_a). \tag{4-13}
 \end{aligned}$$

Será abordado o problema não-paramétrico em primeiro lugar e, posteriormente, o problema paramétrico. As próximas sub-seções mostram as formulações utilizando o modelo proposto neste trabalho e os métodos escolhidos como comparação.

4.1.3.1 Modelagem Proposta

A formulação do problema não-paramétrico é idêntica à formulação 4-6, mas leva em consideração a senóide no lugar da aresta inferior:

$$\begin{aligned}
 &\mathbf{min} F_A[f] \\
 &\text{sujeito a} \\
 f(0, y) &= 0 \rightarrow \sum_{l=1}^{k_a} d_{X,l} f_{X,l}(0) = 0 & \forall y \in [0; 1] \\
 f(1, y) &= 0 \rightarrow \sum_{l=1}^{k_a} d_{X,l} f_{X,l}(1) = 0 & \forall y \in [0; 1] \\
 f(x, \alpha \operatorname{sen}(2\pi kx)) &= 0 \rightarrow \sum_{l=1}^{k_a} d_{Y,l} f_{Y,l}(\alpha \operatorname{sen}(2\pi kx)) = 0 & \forall x \in [0; 1] \\
 f(x, 1) &= 0 \rightarrow \sum_{l=1}^{k_a} d_{Y,l} f_{Y,l}(1) = 0 & \forall x \in [0; 1]. \tag{4-14}
 \end{aligned}$$

O cálculo dos coeficientes em 4-14 é similar à técnica usada em 4-6: monta-se um sistema linear nos coeficientes $d_{X,l}$ e $d_{Y,l}$, $l = 1, 2, \dots, k_a$, e utiliza-se um método de otimização local para minimizar $F_A[f]$ sujeito às restrições lineares estabelecidas pelo referido sistema. No entanto, a restrição envolvendo

a senóide não é imediatamente usável, pois existem, na realidade, infinitas restrições: uma para cada valor particular de x no intervalo $[0; 1]$.

Para contornar essa dificuldade, proponho utilizar uma base ortonormal para o espaço de Hilbert $L^2(\mathbb{R})$ e projetar a restrição $\sum_{l=1}^{k_a} d_{Y,l} f_{Y,l}(\alpha \text{sen}(2\pi kx)) = 0$ em cada uma das funções que compõem a base, para assim obter equações lineares que não dependem da variável x . Em particular, seja $B = \{\psi_{m,n} : m, n = 0, \pm 1, \pm 2, \dots\}$ uma base ortonormal para $L^2(\mathbb{R})$ de modo que toda a função $h \in L^2(\mathbb{R})$ possa ser escrita como $h(x) = \sum_{m,n} \zeta_{m,n} \psi_{m,n}(x)$, onde

$$\zeta_{m,n} = \langle h; \psi_{m,n} \rangle_{\#1} = \int_{-\infty}^{+\infty} h(x) \overline{\psi_{m,n}(x)} dx, \quad (4-15)$$

e $\langle \cdot; \cdot \rangle_{\#1}$ é o produto interno em $L^2(\mathbb{R})$. Em particular, podemos usar a base ortonormal B para escrever

$$\sum_{l=1}^{k_a} d_{Y,l} f_{Y,l}(\alpha \text{sen}(2\pi kx)) = \sum_{m,n} \zeta_{m,n} \psi_{m,n}(x), \quad (4-16)$$

onde

$$\begin{aligned} \zeta_{m,n} &= \int_{-\infty}^{+\infty} \sum_{l=1}^{k_a} d_{Y,l} f_{Y,l}(\alpha \text{sen}(2\pi kx)) \overline{\psi_{m,n}(x)} dx \\ &= \sum_{l=1}^{k_a} d_{Y,l} \int_0^1 f_{Y,l}(\alpha \text{sen}(2\pi kx)) \overline{\psi_{m,n}(x)} dx, \end{aligned} \quad (4-17)$$

pois $f_{Y,l}$ é nula fora do intervalo $[0; 1]$. Finalmente, basta exigir que

$$\begin{aligned} \zeta_{m,n} &= 0 \\ \forall m, n &\in 0, \pm 1, \pm 2, \dots \end{aligned} \quad (4-18)$$

para garantir que a restrição original seja satisfeita. Na prática, o conjunto B deve ser truncado para obtermos um conjunto finito de restrições do tipo 4-18. Se forem usadas Q funções de base no total, teremos que ter $k_a \geq Q + 2$, pois teríamos $Q + 1$ equações em 4-14 na variável y .

Existem várias opções na literatura para o conjunto B . Dentre elas, as exponenciais complexas presentes nas séries de Fourier, os polinômios ortogonais de Legendre e Laguerre (ver [40, 41, 42]) e as funções wavelet. Neste trabalho, optou-se por usar wavelets de Meyer, pelos seguinte motivos:

- os coeficientes $\zeta_{m,n}$ do wavelet de Meyer são sempre reais.
- a série $h(x) = \sum_{m,n} \zeta_{m,n} \psi_{m,n}(x)$ converge “mais rápido” utilizando wavelets do que utilizando polinômios ortogonais ou exponenciais complexas.

Na prática, com wavelets de Meyer será possível usar menos elementos em B , o que implica em menores valores de k_a e, conseqüentemente, espaços de busca menores.

Mais detalhes sobre wavelets e wavelets de Meyer podem ser encontrados em [43, 44, 45].

Para resolver a versão paramétrica do problema, foi usada uma formulação análoga àquela escolhida para a versão retangular do domínio apresentada em 4-11. Mais especificamente, escreve-se $f_a(x, y) = f_X(x)f_Y(y)g_X(a - x)g_Y(a - y)$ e foram aplicadas as restrições como em 4-11. Na aresta senoidal, utiliza-se a mesma abordagem baseada em funções Wavelet de Meyer, já explicada nos parágrafos anteriores.

As próximas subseções mostram a formulação por Método dos Elementos Finitos e Método das Diferenças Finitas.

4.1.3.2

Formulação por Método das Diferenças Finitas

A formulação por Método das Diferenças Finitas parte da definição de uma malha de pontos para representar o conjunto S de maneira discreta. Posteriormente, escreve-se o operador de Laplace aplicado à função f num dado ponto da malha como sendo uma combinação linear dos valores de f em pontos adjacentes. Com isso, ao escrever $-\nabla^2 f = \lambda f$, desenha-se uma equação matricial de autovalores. Este tipo de problema, por sua vez, é resolvido por meio de algoritmos bem conhecidos na área de álgebra computacional.

Para definir a malha de pontos, é conveniente fixar um valor de x e discretizar na direção y , com $\alpha \sin(2\pi kx/a) \leq y \leq 1$ para um valor particular do parâmetro a , utilizando uma quantidade pré-estabelecida de pontos. Em particular, seja G o número de pontos na direção y , fixada a coordenada x . Então, o passo δy é dado por

$$\delta y = \frac{1 - \alpha \sin(2\pi kx/a)}{G - 1}. \quad (4-19)$$

Supondo também G pontos na direção x , o passo δx é dado por

$$\delta x = \frac{1}{G - 1}. \quad (4-20)$$

Logo, a malha é construída variando o valor de x no intervalo $[0; a]$ com passo δx , começando do valor 0 e terminando no valor a . Para cada valor de x , os pontos da malha são definidos discretizando na direção y com passo δy . O total de pontos é igual a G^2 . A figura 7 mostra um exemplo com $G = 15$.

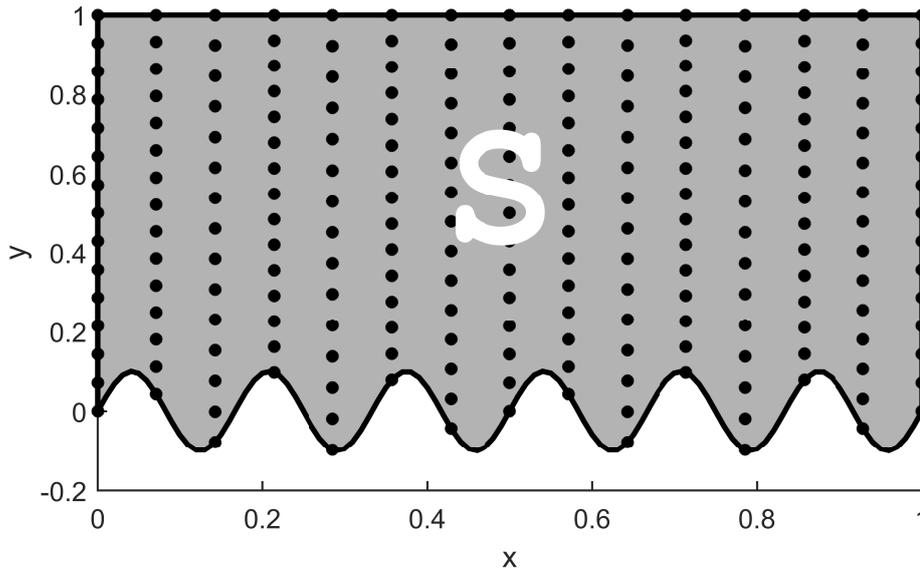


Figura 7: Domínio S com $k = 6$, $a = 1$ e $\alpha = 0.1$, discretizado para o Método Das Diferenças finitas, utilizando 15 pontos por direção (total: 15^2 pontos).

O laplaciano de f no ponto $(x_m; y_n)$ é dado por

$$\begin{aligned} \nabla^2 f(x_m, y_n) = & \frac{f(x_m + \delta x, y_n) - 2f(x_m, y_n) + f(x_m - \delta x, y_n)}{\delta_x^2} \\ & + \frac{f(x_m, y_n + \delta y(x_n)) - 2f(x_m, y_n) + f(x_m, y_n - \delta y(x_n))}{\delta_y(x_n)^2}. \end{aligned} \quad (4-21)$$

A notação $\delta_y(x_n)^2$ é usada para enfatizar que o passo na direção y depende do valor de x .

Ao aplicar a equação 4-21 para cada ponto da malha levando em conta que os pontos no contorno de S são nulos, o problema de autovalores original assume a versão discreta

$$M\mathbf{f} = \lambda\mathbf{f}, \quad (4-22)$$

onde M é uma matriz numérica e \mathbf{f} é um vetor formado pelos valores da função f nos pontos da malha. Este problema pode ser resolvido por algoritmos de álgebra linear computacional, como o Método de Arnoldi [46], implementado em Matlab pela função *eigs* [47].

Embora este método possa parecer simples e apresentar a vantagem de encontrar um conjunto grande de autovalores de uma vez só, ele apresenta uma limitação clara: a presença de uma discretização para representar o domínio. Em particular, a senóide presente no domínio S deve ser discretizada com uma quantidade de pontos muito maior do que a frequência k para que o contorno

senoidal seja representado com acurácia. Deste modo, para valores grandes de k , o valor de G deve ser muito alto. Como a malha aumenta com G^2 , essa discretização facilmente geraria matrizes muito grandes no caso de valores altos de k , o que aumentaria o tempo de execução do Método de Arnoldi para resolver a equação 4-22.

Outro ponto negativo do Método das Diferenças Finitas é a dependência que a discretização do domínio apresenta em relação ao domínio do problema. Mais especificamente, para resolver o problema paramétrico, é necessário repetir o processo de discretização e solução de 4-22 para cada valor particular do parâmetro a . Mais a frente, a seção de resultados irá comparar o tempo de execução do método proposto neste trabalho, cuja solução é analítica e parametrizada por a , com o tempo de execução do Método das Diferenças Finitas, evidenciando o peso de roda-lo para cada valor de a .

4.1.3.3

Formulação por Método dos Elementos Finitos

A formulação por Método dos Elementos Finitos é muito similar àquela apresentada para o Método das Diferenças Finitas. A principal diferença está no tipo de malha: enquanto a malha do Método das Diferenças Finitas é formada por pontos, a malha do Método dos Elementos Finitos é formada por elementos geométricos, tipicamente triângulos ou retângulos. No interior de cada elemento, a função f é modelada por meio de interpolações polinomiais de seus valores nos vértices dos elementos. Opcionalmente, pode-se adicionar mais pontos no contorno de cada elemento, além dos vértices.

Com a presença de uma interpolação no interior dos elementos, o valor de f pode ser aproximado em todos os pontos do domínio, ao contrário do Método das Diferenças Finitas, que só modela a função f para o conjunto finito de pontos da malha. Além disso, a interpolação polinomial simplifica o cálculo das derivadas, pois as derivadas de f passam a ser derivadas de polinômios.

Assim como na modelagem por Método das Diferenças Finitas, a modelagem por Método dos Elementos Finitos também gera um problema matricial de autovalores do tipo

$$M\mathbf{f} = \lambda\mathbf{f}, \quad (4-23)$$

mas, nesse caso, o vetor \mathbf{f} contém os valores de f nos vértices dos elementos e, opcionalmente, em pontos sobre as arestas dos mesmos. Com isso, o sistema 4-23 tende a ser maior e mais complexo do que o sistema 4-22, gerando tempos

de execução maiores do que aqueles típicos do Método Das Diferenças Finitas. Contudo, o Método dos Elementos Finitos tende a ser mais preciso e por isso é o mais usado na indústria.

Apresentar mais detalhes acerca do Método dos Elementos Finitos seria uma tarefa longa e fora de escopo. Ao leitor interessado, recomendo [1, 2, 3]. Neste trabalho em particular, utilizou-se a discretização triangular fornecida pelo *toolbox* de Equações Diferenciais Parciais do Matlab [48] para gerar a malha e o sistema 4-23. A próxima seção irá exibir resultados comparativos entre os métodos citados e o método proposto neste trabalho.

4.1.3.4 Resultados Comparativos

Nesta seção, são apresentados os resultados comparando o modelo proposto, o Método dos Elementos Finitos e o Método das Diferenças Finitas para resolver o problema de autovalores do operador de Laplace com o domínio especificado em 4-12 e 4-13. Em particular, a comparação focará em exibir como a acurácia e o tempo de execução variam com o aumento de complexidade do domínio para cada técnica em questão. Além disso, os experimentos também objetivam demonstrar a vantagem, em tempo de execução, da técnica proposta nesta tese quando se deve fazer uma análise para vários valores de um parâmetro.

Para definir a acurácia de determinação de um autopar aproximado $(\lambda^*; f^*)$, foi utilizado o erro definido como

$$\epsilon = \frac{\|-\nabla^2 f^* - \lambda^* f^*\|}{\|f^*\|}. \quad (4-24)$$

Claramente, $-\nabla^2 f^* = \lambda^* f^*$ se, e somente se, $\epsilon = 0$. Com isso, considera-se que a técnica mais acurada é aquela que gera o menor valor de ϵ .

A tabela 5 mostra os parâmetros utilizados no modelo de PG proposto nesta tese. Além disso, as tabelas 6 e 7 mostram os parâmetros utilizados pelo Método dos Elementos Finitos e pelo Método das Diferenças Finitas.

Em um primeiro experimento, foi analisada a acurácia de cada método para calcular o menor autovalor e sua autofunção, variando a frequência da senóide, k . A malha do Método das Diferenças Finitas e do Método dos Elementos Finitos utilizou 10^2 e 100^2 pontos (são 2 experimentos para cada método). O modelo proposto foi rodado 40 vezes para cada valor de k . A figura 8 apresenta os resultados.

Pelo resultado apresentado na fig. 8, o modelo proposto nesta tese tem perda de acurácia menor do que o Método dos Elementos Finitos e o Método

Tabela 5: Parâmetros do modelo proposto para achar os autopares do operador de Laplace no domínio S .

População	30 indivíduos
Gerações	500 gerações
Árvores por função (k_a)	12
Superposição de árvores	6
Taxa de mutação	fixa em 10%
Número de funções wavelet	9 ($\psi_{0,0}, \psi_{0,\pm 1}, \psi_{\pm 1,0}, \psi_{\pm 1,\pm 1}$)
Tipos de mutação	6 tipos descritos no cap. 3
Tipo de recombinação	sub-árvore
Funções	$\{\sin(t), \cos(t), t^2, t^3, t^4, *, +, -\}$
Terminais	$\{t, \pi, 0, 1, -1\}$
Otimizador Local	<i>BFGS Quasi-Newton</i> [34, 35]
Avaliação	alg. 1 + derivadas e simplificação simbólicas
Profundidade Máxima	4 níveis por árvore

Tabela 6: Parâmetros utilizados pelo Método dos Elementos Finitos para achar os autopares do operador de Laplace no domínio S .

Discretização	Triangular
Interpolação	Polinômios grau 3
Malhador	<i>PDE Toolbox</i> do Matlab [48]
Algoritmo de solução	Método de Arnoldi [47]

Tabela 7: Parâmetros utilizados pelo Método das Diferenças Finitas para achar os autopares do operador de Laplace no domínio S .

Discretização	Malha de Pontos
Interpolação	Não se aplica
Malhador	Conforme descrito na seção 4.1.3.2
Algoritmo de solução	Método de Arnoldi [47]

das Diferenças Finitas. Isso ocorre porque o método proposto é analítico, enquanto os métodos usados como comparação dependem de uma discretização do domínio que fica cada vez menos precisa a medida que a senoide aumenta de frequência. Em particular, na frequência máxima utilizada no experimento ($k = 105$), o erro do Método das Diferenças Finitas com malha de 10^2 pontos chega a ser 3 vezes superior aquele apresentado pela técnica proposta.

Obviamente, aumentar o número de pontos da malha reduz o erro dos métodos baseados em discretização. Contudo, esta possível solução aumentaria o tempo de execução dos métodos, tornando-os possivelmente mais lentos do que o método proposto nesta tese. Por isso, foi realizado um segundo experimento para determinar qual dentre os três métodos atinge um erro pré-estabelecido em menos tempo, para um valor alto de k . A figura 9 exibe os resultados deste experimento.

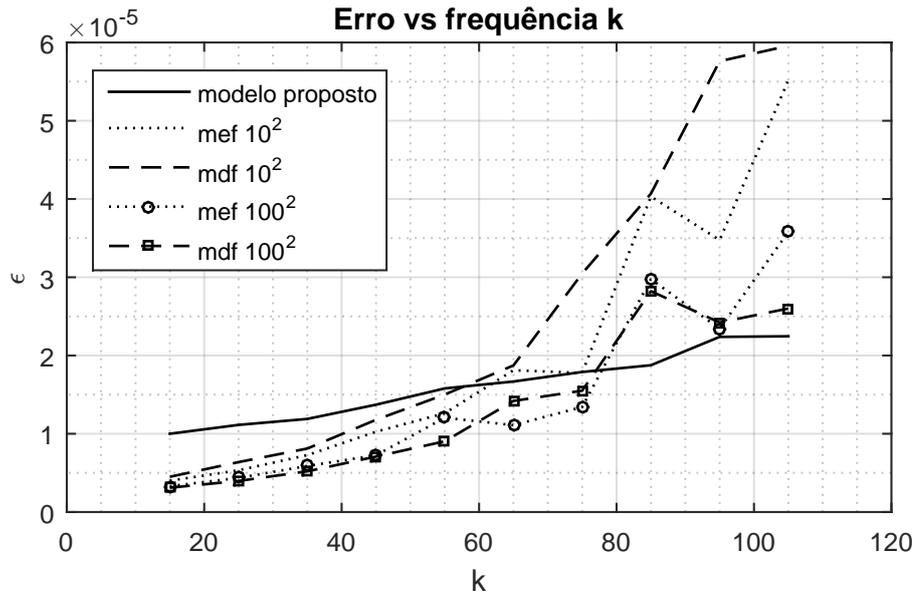


Figura 8: Comparativo de acurácia entre os métodos para valores de k variando entre 15 e 105, com passo de 10.

A figura 9 merece uma explicação mais aprofundada. Em primeiro lugar, todos os experimentos utilizaram $k = 200$, de maneira a tornar o domínio S suficientemente difícil para todos os três métodos. Para obter a curva referente ao modelo proposto, optou-se por parar a evolução da programação genética apenas quando o melhor indivíduo atingir um erro pré-especificado. Dessa forma, a evolução não é interrompida por número de gerações executadas, mas sim por acurácia exigida. Como o algoritmo evolutivo é estocástico, a curva referente ao modelo proposto apresenta o tempo médio de execução deste para atingir a acurácia desejada. Finalmente, as curvas do Método dos Elementos Finitos e Método das Diferenças Finitas foram obtidas, para cada cada erro pré-estabelecido, através de sucessivos refinamentos da malha discreta até conseguir atingir um erro menor ou igual àquele pré-estabelecido.

Analisando o gráfico 9, observa-se que, para valores de erro maiores, as técnicas clássicas têm tempo de execução menores, pois trabalham com malhas relativamente pequenas. Por outro lado, a medida em que os valores de erro-alvo são reduzidos, a necessidade de refinamentos sucessivos na malha eleva consideravelmente o tempo de execução do Método dos Elementos Finitos e do Método das Diferenças Finitas. Assim, para valores de erro inferiores a 10^{-5} , a técnica proposta tem tempo de execução médio inferior àquele apresentado pelas técnicas clássicas, pois não necessita de discretização refinada para representar de maneira acurada a aresta senoidal de alta frequência. Este comportamento evidencia uma vantagem de se trabalhar com um método

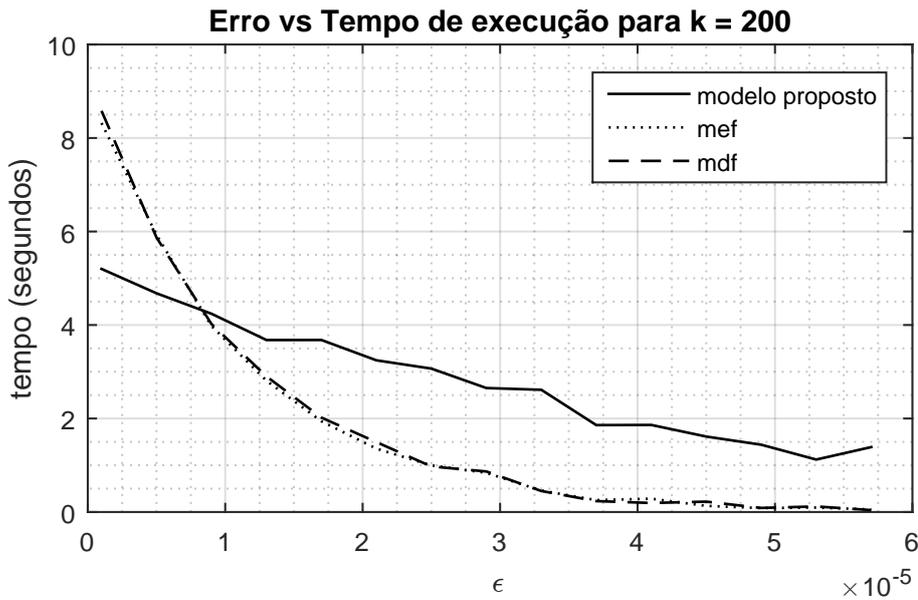


Figura 9: Comparação entre o tempo de execução dos métodos para atingir um erro menor ou igual ao um valor pré-estabelecido. O experimento foi executado com $k = 200$.

puramente analítico: como não há discretização de domínio, os domínios complexos não aumentam significativamente o tempo de execução a medida que os requisitos de acurácia tornam-se mais rigorosos.

Para concluir esta seção, é importante fazer uma análise da solução da versão paramétrica do problema, definido em 4-13. Neste caso, o interesse é analisar a acurácia da solução paramétrica apresentada, comparando-a, para cada valor do parâmetro a , com aquela obtida pelo Método dos Elementos Finitos e pelo Método das Diferenças Finitas para o mesmo valor do parâmetro. Neste caso, como os métodos clássicos não conseguem entregar uma solução parametrizada, é necessário executa-los para cada valor particular do parâmetro a .

De modo a garantir uma comparação justa, foi escolhido um tempo de execução fixo de 2s para cada execução dos modelos de Método dos Elementos Finitos e Método das Diferenças Finitas. Em outras palavras, para cada valor particular do parâmetro a , foi escolhido o grid mais refinado que ainda implicaria em um tempo de execução inferior a 2s. O experimento foi conduzido para 5 valores distintos de a e, por isso, foi permitido que o método proposto evoluísse até que o tempo de execução fosse igual ao tempo de 5 execuções da técnica clássica, isto é, 10s. Portanto, foram feitas 5 execuções do Método das Diferenças Finitas, cada uma com 2s de duração, 5 execuções do Método dos Elementos Finitos, cada uma com 2s de duração, e 1 execução do modelo proposto, com 10s de duração. As tabelas 8 e 9 apresentam o erro obtido por

cada método para cada um dos 5 valores de a escolhidos.

Tabela 8: Acurácia obtida resolvendo o problema paramétrico para achar o primeiro autopar do operador de Laplace no domínio S , utilizando $k = 200$.

Valor do parâmetro (a)	Método Proposto	MEF	MDF
0.5	2.03×10^{-5}	1.87×10^{-5}	1.91×10^{-5}
0.75	1.91×10^{-5}	1.82×10^{-5}	1.97×10^{-5}
1.00	2.12×10^{-5}	1.88×10^{-5}	2.02×10^{-5}
1.25	1.98×10^{-5}	1.84×10^{-5}	1.95×10^{-5}
1.5	1.91×10^{-5}	1.91×10^{-5}	2.01×10^{-5}

Tabela 9: Acurácia obtida resolvendo o problema paramétrico para achar o segundo autopar do operador de Laplace no domínio S , utilizando $k = 200$.

Valor do parâmetro (a)	Método Proposto	MEF	MDF
0.5	2.12×10^{-5}	1.93×10^{-5}	2.11×10^{-5}
0.75	2.01×10^{-5}	1.90×10^{-5}	2.01×10^{-5}
1.00	2.07×10^{-5}	1.91×10^{-5}	2.08×10^{-5}
1.25	1.96×10^{-5}	1.89×10^{-5}	1.99×10^{-5}
1.5	1.99×10^{-5}	1.98×10^{-5}	1.95×10^{-5}

Observa-se pelas tabelas 8 e 9 que a solução analítica parametrizada é capaz de modelar os dois primeiros autopares do operador de Laplace com acurácia muito semelhante àquela obtida pelos métodos clássicos. Consequentemente, é possível reutilizar a autofunção parametrizada por a para qualquer valor particular de a dentro da faixa $[0.5; 1.5]$, obtendo instantaneamente a estimativa do autopar com acurácia semelhante a dos métodos clássicos. Em aplicações que necessitem determinar o autovalor para muitos cenários de a , o modelo proposto torna-se, portanto, mais eficiente, pois necessita de apenas uma execução, contra as muitas execuções que seriam necessárias utilizando uma técnica clássica.

Sendo assim, o modelo de solução analítica baseado em PG mostrou-se capaz de encontrar os autovalores do operador de Laplace em domínios complexos, com acurácia melhor que as técnicas clássicas supondo domínio suficientemente complexo. Além disso, a técnica proposta também foi capaz de resolver o problema paramétrico, entregando uma solução analítica que recebe o parâmetro como um dos argumentos. Esta função apresentou acurácia semelhante aos métodos clássicos, quando executados para um valor específico do parâmetro. Contudo, por apresentar uma solução paramétrica, o método proposto tende a apresentar tempo de execução inferior às técnicas clássicas, uma vez que sua solução pode ser reutilizada infinitamente, enquanto os métodos clássicos necessitam ser re-executados para cada valor particular do

parâmetro. Na próxima seção, será feita uma apresentação de outro operador auto-adjunto de interesse na Física: o oscilador harmônico quântico.

4.2

Oscilador Harmônico Quântico (OHQ)

O oscilador harmônico quântico (OHQ) [49] é um sistema físico análogo ao oscilador harmônico clássico derivado de sistemas massa-mola. Na Física Quântica e áreas de aplicação, é extremamente importante pois é um sistema relativamente simples e que serve de modelo para resolver problemas mais complexos.

O problema de autovalores é dado, de maneira simplificada, por

$$\begin{aligned}
 A\Psi &= \lambda\Psi \\
 A\Psi &= -\Psi_{[xx]} + \alpha^2 x^2 \Psi \\
 \lambda &= \frac{8\pi^2 m E}{h^2} \\
 \alpha &= \frac{2\pi m \omega}{h} \\
 \Psi(x) &\in L^2(\mathbb{R})
 \end{aligned} \tag{4-25}$$

onde h é a constante de Plank, m é a massa da partícula, ω é a frequência angular de oscilação e E é a energia, que é uma constante positiva a ser determinada a partir do autovalor λ . Para simplificar, assume-se que a frequência ω é escolhida de modo que $\alpha = 1$. Neste caso, é possível mostrar que a solução exata não normalizada é dada por

$$\begin{aligned}
 \Psi_n(x) &= H_n(x) e^{-\frac{x^2}{2}}, \\
 \lambda_n &= \frac{4\pi m \omega (n + 1/2)}{h}, \\
 n &= 0, 1, 2, 3, \dots, \\
 x &\in \mathbb{R}
 \end{aligned} \tag{4-26}$$

onde $H_n(x)$ é o polinômio de Hermite físico de ordem n , definido por

$$\begin{aligned}
 H_n(x) &= (-1)^n e^{x^2} \frac{\partial^n}{\partial x^n} e^{-x^2}, \\
 n &= 0, 1, 2, 3, \dots, \\
 x &\in \mathbb{R}.
 \end{aligned} \tag{4-27}$$

Os primeiros polinômios de Hermite são

$$\begin{aligned} H_0(x) &= 1, \\ H_1(x) &= 2x, \\ H_2(x) &= 4x^2 - 2, \\ H_3(x) &= 8x^3 - 12x. \end{aligned} \tag{4-28}$$

O operador A definido em 4-25 é auto-adjunto não-negativo no produto interno

$$\begin{aligned} \langle f; g \rangle &= \int_0^{+\infty} f \bar{g} dx, \\ f, g &\in L^2(\mathbb{R}). \end{aligned} \tag{4-29}$$

O problema não paramétrico foi resolvido utilizando o conjunto de parâmetros mostrados na tabela 10. Neste caso, note que exponenciais e polinômios de Hermite foram acrescentados para permitir que o modelo encontre a solução exata.

Tabela 10: Modelo proposto para achar os autopares do oscilador harmônico quântico.

População	500 indivíduos
Gerações	2000 gerações
Núm. Árvores (k_a)	de 1 a 5 (vários experimentos)
Taxa de mutação	fixa em 10%
Tipos de mutação	6 tipos descritos no cap. 3
Tipo de recombinação	sub-árvore
Funções	$\{e^t, t^2, H_1(t), *, +, /\}$
Terminais	$\{x, 0, 1, -1\}$
Otimizador Local	<i>BFGS Quasi-Newton</i> [34, 35]
Avaliação	alg. 1 + derivadas e simplificação simbólicas
Profundidade Máxima	4 níveis por árvore

É importante mencionar que não é necessário colocar outras funções de Hermite, porque elas se relacionam recursivamente por

$$\begin{aligned} H_{n+1}(x) &= 2xH_n(x) - 2nH_{n-1}(x), \\ n &= 1, 2, 3, \dots, \\ x &\in \mathbb{R}, \end{aligned} \tag{4-30}$$

e $H_0(t) = 1$, que já está presente nos terminais. Outro ponto importante é que, como as funções de Hermite são polinômios, a sua presença na lista de funções é opcional, dado que x está no conjunto de terminais e as operações

de multiplicação e soma estão presentes na lista de funções. Contudo, isso certamente iria forçar o modelo a encontrar soluções extensas e difíceis de ler. Pelo mesmo motivo, acrescentou-se $f(t) = t^2$ na lista de funções.

Tabela 11: Resultados para a primeira autofunção do oscilador harmônico quântico.

Num. Árvores	Ψ_0	$b\Psi_0, b \neq 1$	Outra
1	21	12	7
2	17	13	10
3	19	12	9
4	15	10	15
5	10	9	21

A tabela 11 apresenta os resultados após 40 rodadas para encontrar a autofunção relacionada ao menor autovalor. Novamente, observa-se que o modelo encontra a solução exata em mais da metade das rodadas (exceto no caso de 5 árvores, onde outra função foi encontrada 21 vezes das 40 rodadas). Isso mostra que o modelo encontra com sucesso a função exata, especialmente nos casos em que o espaço de busca é menor. Para confirmar este resultado, outro experimento foi realizado, agora com 4000 gerações, e os resultados deste foram melhores, conforme na tabela 12.

Tabela 12: Resultados para a primeira autofunção do oscilador harmônico quântico – 4000 gerações.

Num. Árvores	Ψ_0	$b\Psi_0, b \neq 1$	Outra
1	32	4	4
2	33	4	3
3	30	6	4
4	28	10	2
5	23	13	4

Como mostra a tabela 12, dobrar o número de gerações faz com que o modelo encontre a autofunção relativa ao menor autovalor na maioria das rodadas, mesmo com 5 árvores. Isto confirma que o modelo converge para a resposta se o número de gerações for suficientemente grande.

Numa segunda análise, definiu-se o erro $\epsilon = \|A\Psi - \lambda\Psi\|/\|\Psi\|$ e o modelo proposto foi executado para encontrar o menor autovalor até que $\epsilon < 10^{-4}$. Foram feitas novamente 40 rodadas e observou-se quantas gerações foram necessárias para obter o referido erro. A tabela 13 apresenta este resultado sob forma de histograma.

Os resultados mostram que, para um número pequeno de árvores, a maior parte dos experimentos converge com 2500 a 3500 gerações. Além disso,

Tabela 13: Resultados para a primeira autofunção do OHQ – número de rodadas que convergiram com erro $< 10^{-4}$ (células) para cada faixa de gerações (colunas).

Num. Árvores	< 1500	de 1500 a 2500	de 2500 a 3500	> 3500
1	6	29	5	0
2	2	30	4	4
3	0	33	1	6
4	0	24	9	7
5	0	17	11	12

com um número alto de árvores, a convergência tende a consumir mais de 3500 gerações, devido ao maior espaço de busca. Esses resultados mostram novamente que o modelo pode ser usado com robustez para encontrar a autofunção com um erro pré-especificado. Naturalmente, o número total de gerações depende da inicialização (população inicial).

Finalmente, utilizou-se o modelo para encontrar a autofunção referente ao segundo menor autovalor ($n = 1$). Novamente, o modelo encontrou o autovalor exato na maior parte dos casos. Este caso, no entanto, é semelhante em termos de complexidade ao caso do menor autovalor pois todas as funções necessárias estão imediatamente presentes no conjunto de funções (tabela 10). No entanto, para encontrar autofunções de ordem maior, o modelo precisa deduzir uma combinação de funções de Hermite de ordem menor, no caso $H_1(t)$ e $H_0(t)$. A tabela 14 mostra os resultados assumindo evolução até encontrar a autofunção com erro menor que 10^{-4} .

Tabela 14: Resultados para a 3a. menor autofunção do OHQ – número de rodadas que convergiram com erro $< 10^{-4}$ (células) para cada faixa de gerações (colunas).

Num. Árvores	< 1500	de 1500 a 2500	de 2500 a 3500	> 3500
1	1	11	17	11
2	12	23	5	0
3	7	31	2	0
4	6	22	3	9
5	2	20	11	7

O resultado da tabela 14 mostra claramente que, diferentemente dos experimentos anteriores, o modelo com 2 árvores comportou-se melhor, pois houve a maior taxa de convergência com menos de 1500 gerações. Isso se deve ao fato de que a solução exata é na realidade um somatório de duas parcelas, pois $H_2(x) = 2xH_1(x) - 2$ e a autofunção exata é dada por 4-26 com $n = 2$. Em casos deste tipo, onde a solução precisa ser expressa como uma combinação linear de funções, o modelo multi-árvore desempenha

consideravelmente melhor, como foi o caso deste experimento utilizando 2, 3 e 4 árvores.

Por fim, foi realizado um experimento para entender o comportamento da técnica proposta quando as funções que compõem a solução exata não estão presentes no conjunto de funções fornecidas. Para isso, tentou-se encontrar a terceira autofunção ($n = 2$) sem o conhecimento das funções de Hermite. A tabela 15 mostra os parâmetros utilizados pelo modelo proposto.

Tabela 15: Modelo proposto para achar o terceiro autoper do oscilador harmônico quântico (sem utilizar funções de Hermite).

População	500 indivíduos
Gerações	2000 gerações
Núm. Árvores (k_a)	de 1 a 5 (vários experimentos)
Taxa de mutação	fixa em 10%
Tipos de mutação	6 tipos descritos no cap. 3
Tipo de recombinação	sub-árvore
Funções	$\{e^t, t^2, t^3, \text{sen}(t), \text{cos}(t), *, +, /\}$
Terminais	$\{x, 0, 1, -1\}$
Otimizador Local	<i>BFGS Quasi-Newton</i> [34, 35]
Avaliação	alg. 1 + derivadas e simplificação simbólicas
Profundidade Máxima	5 níveis por árvore

Observe na tabela 15 que as funções t^3 , $\text{sen}(t)$ e $\text{cos}(t)$ foram colocadas no lugar das funções de Hermite, sendo que as funções trigonométricas foram usadas para “confundir” a técnica, uma vez que elas não são necessárias na resposta final.

Neste novo experimento, a técnica proposta foi executada por 40 vezes e os resultados podem ser observados na tabela 16.

Tabela 16: Resultados para a terceira autofunção do OHQ – número de rodadas que convergiram com erro $< 10^{-4}$ (células) para cada faixa de gerações (colunas), experimento sem funções de Hermite.

Num. Árvores	< 1500	de 1500 a 2500	de 2500 a 3500	> 3500
1	2	18	9	11
2	0	27	4	9
3	0	30	1	9
4	0	25	3	12
5	0	21	7	12

Pelos resultados apresentados, observa-se que a técnica teve mais dificuldade de encontrar autofunções com erro menor do que 10^{-4} do que no experimento com funções de Hermite, como já era esperado. Contudo, ainda assim o modelo proposto foi capaz de solucionar o problema de autovalores, o

que mostra que, mesmo que o usuário da técnica não tenha conhecimento do tipo de função que deve colocar na base de funções, a técnica ainda é capaz de encontrar autofunções satisfatórias.

Por fim, é interessante analisar a expressão analítica da melhor autofunção encontrada e compara-la com a resposta teórica. A resposta teórica é dada pela equação 4-26 para $n = 2$:

$$\Psi_2(x) = (4x^2 - 2)e^{-x^2/2}. \quad (4-31)$$

A melhor autofunção encontrada apresentou erro de 1.6008×10^{-5} e sua expressão, após simplificação simbólica, é dada por

$$\tilde{\Psi}_2(x) = (4,00001257x^2 - 0,00000374x - 2,000000107)e^{-0,49999841x^2}. \quad (4-32)$$

Portanto, o resultado mostra que a técnica proposta encontrou uma expressão algébrica muito próxima ao esperado, a menos a parcela $-0,00000374x$, que na prática influencia muito pouco no valor numérico de $\tilde{\Psi}_2(x)$. Sendo assim, concluo que o método proposto foi capaz de encontrar as autofunções do oscilador harmônico quântico, mesmo quando as funções necessárias para tal não estão diretamente presentes na base de funções fornecidas ao algoritmo de PG.

4.3

Conclusões do Capítulo

O experimento do OHQ, juntamente com o experimento do Operador de Laplace, constituem estudos de caso simples cuja solução exata é conhecida e, portanto, são bons candidatos para fazer prova de conceito. Os resultados deste capítulo mostram que o modelo proposto tem uma boa robustez, e de fato converge para a solução exata com um número suficientemente grande de gerações.

Além disso, o capítulo também mostrou uma versão do problema do operador de Laplace que não apresenta solução exata, pois uma das arestas do domínio retangular original foi substituída por uma senoide. Assim, como forma de comparação, foram empregadas as técnicas clássicas, conhecidas como o Método dos Elementos Finitos e o Método das Diferenças Finitas. Os resultados mostraram que a técnica proposta conseguiu resolver o problema de autovalores com acurácia e tempo de execução melhores que aqueles apresentados pelas técnicas clássicas, sobretudo quando a senoide que delimita o domínio tem alta frequência. Isso mostra que a técnica proposta, por não depender de discretização do domínio, apresenta superioridade em casos cujo

domínio é complexo quando comparada às técnicas baseadas em discretização.

Ainda neste capítulo, também foi apresentado o resultado de solução do problema de Laplace parametrizado pelo tamanho da aresta do domínio retangular. Neste caso, a técnica proposta foi executada para gerar autofunções parametrizadas pelo tamanho da aresta do domínio. Os resultados mostraram que a versão paramétrica tem acurácia semelhante àquela obtida pelo Método dos Elementos Finitos e pelo Método das Diferenças Finitas, o que por si só já representa uma vantagem em relação às técnicas clássicas, uma vez que tal solução pode ser reutilizada para qualquer valor particular do parâmetro, sem necessidade de re-execução da técnica. No caso das técnicas clássicas, que não são paramétricas, é necessário discretizar o domínio e re-executar a técnica clássica para cada valor particular do parâmetro, o que implica em altos tempos de execução acumulados quando necessita-se de uma análise que depende de muitos cenários.

Resta, portanto, mostrar como o modelo se comporta com problemas mais complexos, nos quais a solução exata é desconhecida e o domínio é complexo. Para isso, o próximo capítulo exibirá um estudo de caso baseado em poços de petróleo, com interesse particular em prever os modos de propagação eletromagnética dentro do tubo de produção de um poço.

5

Estudos de Caso: Operadores Não Auto-Adjuntos

5.1

Descrição Preliminar do Problema

Assim como o capítulo 4 este capítulo foca em responder à pergunta “quais são os resultados obtidos?”. No entanto, o capítulo dedica-se a apresentar um estudo de caso prático, no qual o problema é vetorial e leva a operadores não normais.

A motivação deste capítulo é o problema de telemetria sem fio de poços de petróleo em plataformas *offshore*. Neste problema, deseja-se instalar uma rede de sensores operando sem fio que permita comunicação constante e robusta com a plataforma, e em tempo real. Assim, engenheiros em plataforma podem monitorar condições dos poços em vários pontos e tomar decisões de acionamento de válvulas, tratamento químico e reparo de poço com mais acurácia.

Alguns trabalhos estão presentes na literatura para modelar a propagação eletromagnética em reservatórios de petróleo. Na atual literatura, os principais trabalhos focam principalmente em desenvolver modelos assumindo propagação pela formação rochosa vizinha ao poço [50, 51, 52, 53, 54, 55]. Devido à presença de diversas camadas de rocha diferentes, o problema resume-se à radiação de ondas eletromagnéticas em meios não homogêneos divididos em camadas. Os resultados da literatura mostram que há grande atenuação na formação rochosa e, por isso, é necessário utilizar frequências de propagação da ordem de apenas algumas centenas de Hertz, o que reduziria muito a taxa de transmissão de um sistema digital.

Mais recentemente, alguns trabalhos foram publicados analisando a propagação de ondas eletromagnéticas no interior da região anular dos poços, em oposição à propagação pela formação rochosa [52, 55]. Em tal região, o sinal propaga-se como num cabo coaxial com perdas devido às características do fluido presente nesta parte do poço. Além disso, a presença de equipamentos geralmente metálicos dificulta a propagação no interior da região anular, pois aumenta consideravelmente a atenuação do sinal. Por estas razões, a propagação em modo transversal eletromagnético na região anular também

se limita a frequências muito baixas.

Dado este cenário, o presente capítulo procura contribuir apresentando um modelo de propagação eletromagnética no interior da coluna de produção, dentro da qual o óleo flui. Este modelo de propagação utiliza o modelo de solução de problemas de autovalores proposto no capítulo 3 para resolver o problema de autovalores resultante da busca de modos no interior do duto de produção.

A região no interior do duto de produção se caracteriza por não homogeneidades devido à variação da condutividade do óleo ao longo do comprimento do poço e devido à presença de incrustações salinas na parede da coluna de produção. Portanto, o modelo desenvolvido utiliza conceitos de guias de onda não homogêneos e é capaz de calcular os modos híbridos de propagação.

A próxima seção formula matematicamente o problema e apresenta um modelo semi-analítico que foi desenvolvido como ferramenta comparativa. Na sequência, o capítulo apresenta o modelo proposto de solução baseado em programação genética (PG) para este problema. Finalmente, os resultados são apresentados, comparando o modelo semi-analítico com o modelo proposto baseado em PG. Esta comparação inclui experimentos paramétricos e não-paramétricos, considerando duas faixas de frequência de onda. O capítulo conclui fazendo uma análise do desempenho do modelo proposto sob a ótica de acurácia e tempo de execução, indicando as vantagens do modelo proposto em relação à técnica clássica geralmente presente na literatura específica deste problema.

5.2

Formulação Matemática

5.2.1

Premissas deste modelo

A coluna de produção é o duto (ou conjunto de dutos) dentro do qual o óleo flui. É construída a partir de seções de aço, que podem ter diâmetros distintos para garantir determinada pressão ou vazão. No interior da coluna, o petróleo flui obedecendo vários tipos de regime multifásico, com uma mistura de óleo, água salina e gás natural.

A figura 10 dá uma visão geral e simplificada da estrutura geométrica do poço, com a formação rochosa, o revestimento (em cinza), a cimentação de revestimento (com um padrão em “x”), a coluna de produção (em preto), o reservatório de petróleo, e os *packers* de revestimento (linhas grossas cinza horizontais). No caso particular desta figura, a coluna de produção é um

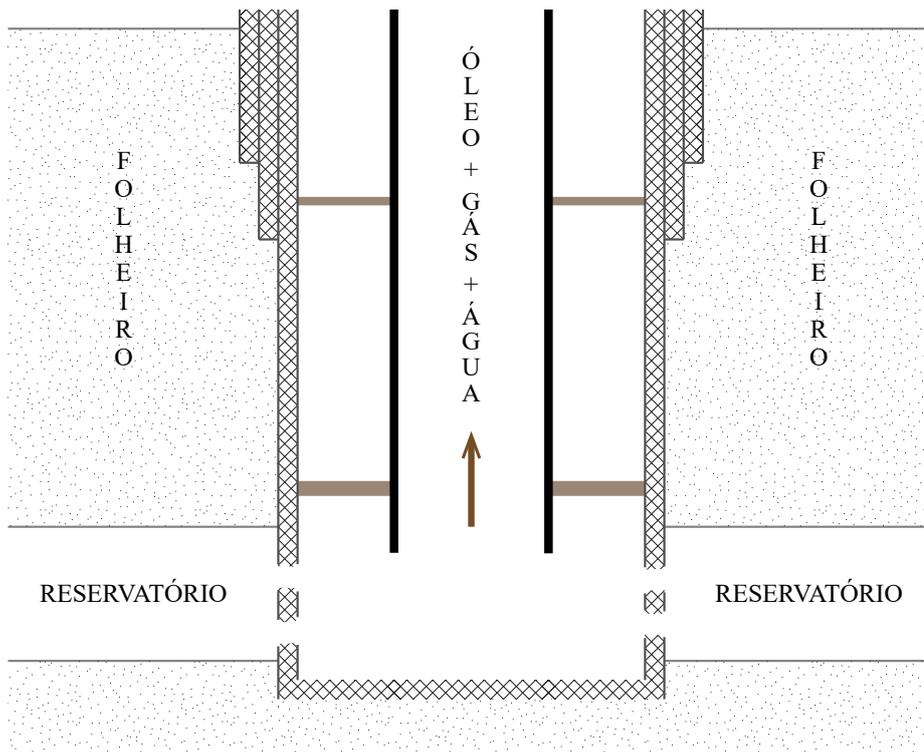


Figura 10: Corte de um poço.

cilindro com diâmetro constante, mas é possível que o diâmetro da coluna varie com a profundidade do poço. Os pequenos furos presentes no revestimento próximos ao reservatório são conhecidos como canhoneamentos e permitem a passagem de óleo para o interior do duto.

Neste trabalho, estamos interessados em modelar a propagação eletromagnética no interior do duto de produção, isto é, na região que circula uma mistura de óleo, gás e água (vide figura 10). O modelo desenvolvido assume algumas simplificações para tornar o problema tratável, esperando, contudo, não sacrificar a acurácia da solução.

A primeira simplificação já aparece na geometria da própria coluna de produção. Tipicamente, a coluna de produção tem formato cilíndrico circular, mas o raio pode variar ao longo do poço a fim de acomodar variações de pressão ou vazão no óleo. Contudo, este trabalho assume que a coluna de produção é um cilindro de base circular e diâmetro constante.

Outras simplificações envolvem as características elétricas dos materiais envolvidos. Os revestimentos e a coluna de produção são feitos com ligas de aço de grande condutividade elétrica. Por isso, o presente trabalho assume que há pouco vazamento de sinal para fora da coluna de produção, embora ainda seja possível ter perdas por efeito Joule. Da mesma forma, este trabalho assume que os revestimentos metálicos têm condutividade igual à do tubo de aço, o que não deve causar perdas de acurácia consideráveis, já que o sinal é

guiado pela coluna de produção, e não pelos revestimentos. Além disso, este trabalho também assume que os *packers* de vedação são metálicos e formados por condutores semelhantes aos do tubo de produção.

O petróleo no interior da coluna de produção é uma mistura de água salina, óleo e gás natural e apresenta duas fases, uma líquida (água + óleo) e uma gasosa. Dependendo do regime de escoamento, algumas disposições geométricas de óleo e gás são comuns no interior da coluna de produção. No regime de bolha, o tubo de produção é completamente preenchido pela mistura de óleo com água e por pequenas bolhas de gás espaçadas de maneira uniforme no interior do líquido. Por outro lado, no regime anular, a fase líquida de óleo e água escoam colada à borda do tubo de produção, enquanto a fase gasosa escoam no centro, de maneira contínua. A quantidade de gás é representada por uma constante, a relação gás-óleo (*GOR*), que é definida como a razão entre a vazão de gás e a vazão de líquidos. Como o gás natural tem condutividade desprezível, quanto maior a razão gás-óleo, menores devem ser as perdas de propagação. O regime de interesse para esta tese é o regime anular, conforme descrito em [56].

Outra característica importante do meio de propagação é a presença de incrustações salinas na parede interior do tubo de produção [57]. Em particular, este trabalho aborda o caso mais comum, no qual CaCO_3 em estado sólido aparece incrustado na parede da coluna de produção, podendo ocupar até 25% do diâmetro interno da coluna. Como a condutividade elétrica do CaCO_3 em estado sólido é desprezível, acredita-se que quanto maior a incrustação, menores serão as perdas de propagação. A tabela 17 resume as premissas feitas para modelar a propagação no interior da coluna de produção.

Tabela 17: Premissas da modelagem da coluna de produção

Premissa	Descrição
Geometria cilíndrica	A coluna apresenta geometria cilíndrica de diâmetro constante.
Coluna de produção	Aço com condutividade finita e elevada, pouco vazamento de sinal.
Revestimentos	Idem à coluna de produção.
Petróleo	Mistura bifásica de óleo com água salina (fase líquida) e gás natural (fase gasosa).
Incrustações	Há CaCO_3 sólido na parede interior da coluna, ocupando até 25% do diâmetro.
Escoamento	Anular, com gás no centro e um anel de óleo na extremidade da coluna.

A figura 11 mostra o regime anular de escoamento, contendo gás na parte central, depois de maneira concêntrica temos, nesta ordem, óleo, incrustação,

duto de aço, revestimento e rocha. Para simplificar o problema, a modelagem assume que a característica anular do gás se mantém no fundo do poço, conforme mostra a figura 11.

Do ponto de vista eletromagnético, o duto de produção se comporta como um guia de ondas infinito preenchido com um dielétrico que tem perdas e não é homogêneo na direção radial. Como a condutividade do óleo varia com a temperatura e esta varia com a profundidade, conclui-se que o meio de propagação também é não-homogêneo na direção de propagação.

5.2.2 Dedução do modelo eletromagnético

Neste trabalho, estamos interessados em encontrar os modos de propagação no duto de produção, assumindo simetria cilíndrica. A formulação deste problema utiliza equações de Maxwell e teoria de Eletromagnetismo clássico.

Considere as quatro equações de Maxwell conforme definidas abaixo,

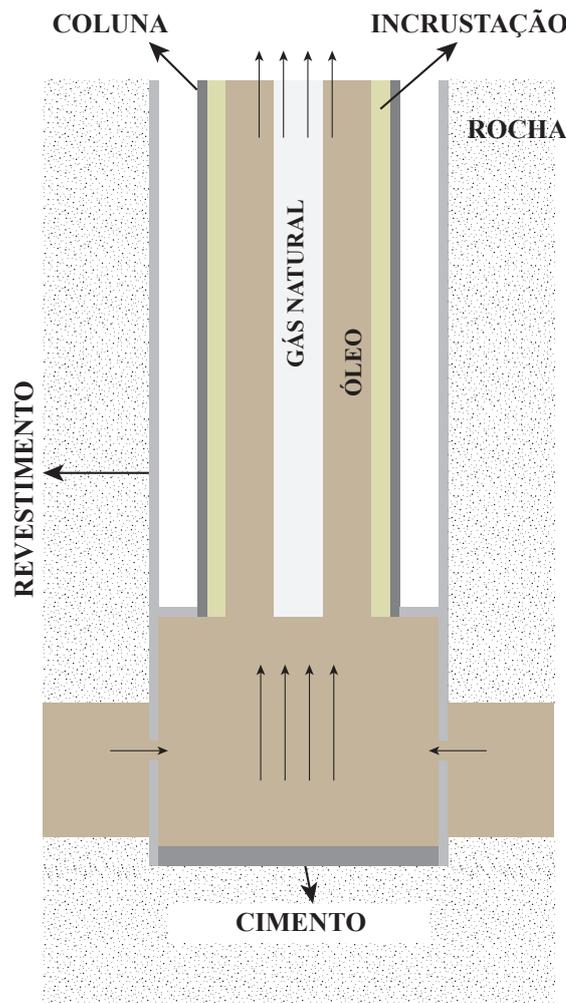


Figura 11: Modelo de coluna de produção com escoamento em regime anular.

onde $i = \sqrt{-1}$, válidas para um meio homogêneo, linear e isotrópico e sem fontes.

$$\nabla \times \mathbf{E} = -i\omega\mu\mathbf{H} \qquad \nabla \times \mathbf{H} = (\sigma + i\omega\epsilon)\mathbf{E} \qquad (5-1)$$

$$\nabla \cdot \mathbf{E} = 0 \qquad \nabla \cdot \mathbf{H} = 0. \qquad (5-2)$$

Nas equações 5-1 e 5-2, \mathbf{E} e \mathbf{H} são o campo elétrico e o campo magnético, respectivamente. Além disso, ω é a frequência angular, μ é a permeabilidade magnética, ϵ é a permissividade elétrica e σ é a condutividade elétrica. O operadores rotacional e divergente são representados pelas notações tradicionais. Todas as equações harmônicas no tempo foram obtidas a partir das equações de Maxwell no tempo usando um par de transformada de Fourier:

$$f(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{f}(t)e^{-i\omega t} dt \qquad (5-3)$$

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(\omega)e^{+i\omega t} d\omega, \qquad (5-4)$$

onde \hat{f} representa o campo variante no tempo e f representa o campo harmônico no tempo, dependente de ω .

A partir de 5-1, deduz-se a equação de onda vetorial, com F representando o campo elétrico ou o campo magnético:

$$\nabla \times \nabla \times \mathbf{F} = \kappa^2 \mathbf{F}, \qquad (5-5)$$

onde

$$\kappa^2 = -\tau(\sigma + i\omega\epsilon) \qquad (5-6)$$

$$\tau = i\omega\mu \qquad (5-7)$$

é a constante de onda do meio.

A partir de agora, usaremos o sistema cilíndrico de coordenadas (ρ, θ, z) com vetores diretores unitários \mathbf{a}_ρ , \mathbf{a}_θ e \mathbf{a}_z . Em [58], os autores mostraram que as soluções da equação de onda 5-5 que também satisfazem 5-2 são:

$$\mathbf{M}_1 = \nabla \times (\mathbf{a}_z u) \qquad (5-8)$$

$$\mathbf{N}_2 = \kappa^{-1} \nabla \times \nabla \times (\mathbf{a}_z v), \qquad (5-9)$$

onde as funções escalares u e v devem satisfazer a equação escalar de Helmholtz:

$$\nabla^2 u + \kappa^2 u = 0 \quad (5-10)$$

$$\nabla^2 v + \kappa^2 v = 0. \quad (5-11)$$

A partir de 5-5, 5-8 e 5-9, escreve-se os campos elétrico e magnético como:

$$\mathbf{E} = q_1 \mathbf{E}_{\text{tm}} + q_2 \mathbf{E}_{\text{te}} \quad \mathbf{H} = q_1 \mathbf{H}_{\text{tm}} + q_2 \mathbf{H}_{\text{te}}, \quad (5-12)$$

onde,

$$\mathbf{E}_{\text{te}} = \mathbf{M}_1 \quad \mathbf{E}_{\text{tm}} = \mathbf{N}_2 \quad (5-13)$$

$$\mathbf{H}_{\text{te}} = -\tau^{-1} \nabla \times \mathbf{M}_1 \quad \mathbf{H}_{\text{tm}} = -\tau^{-1} \nabla \times \mathbf{N}_2. \quad (5-14)$$

As equações em 5-14 referentes ao campo magnético ficam mais simples se definirmos $\mathbf{M}_2 = \nabla \times (\mathbf{a}_z v)$ e $\mathbf{N}_1 = \kappa^{-1} \nabla \times (\mathbf{M}_1)$. Note que $\kappa^{-1} \nabla \times \mathbf{M}_2 = \mathbf{N}_2$ e, como $\nabla \times \nabla \times \mathbf{N}_2 = \kappa^2 \mathbf{N}_2$, segue que $\nabla \times \nabla \times \mathbf{M}_2 = \kappa^2 \mathbf{M}_2$, donde $\nabla \times \mathbf{N}_2 = \kappa^{-1} \nabla \times \nabla \times \mathbf{M}_2 = \kappa \mathbf{M}_2$. Similarmente, $\nabla \times \mathbf{M}_1 = \kappa \mathbf{N}_1$. Assim, as equações 5-14 resumem-se a:

$$\mathbf{H}_{\text{te}} = -\kappa \tau^{-1} \mathbf{N}_1 \quad \mathbf{H}_{\text{tm}} = -\kappa \tau^{-1} \mathbf{M}_2. \quad (5-15)$$

Após expandir, o campo elétrico \mathbf{E}_{te} fica:

$$\mathbf{E}_{\text{te}} = \rho^{-1} u_{[\theta]} \mathbf{a}_\rho - u_{[\rho]} \mathbf{a}_\theta. \quad (5-16)$$

onde o subscripto $[\]$ representa derivadas.

Utilizando 5-8, 5-9, 5-13 e a expressão do laplaciano em coordenadas cilíndricas, temos:

$$\mathbf{E}_{\text{tm}} = \kappa^{-1} \left\{ v_{[z\rho]} \mathbf{a}_\rho + \rho^{-1} v_{[z\theta]} \mathbf{a}_\theta + (v_{[zz]} + \kappa^2 v) \mathbf{a}_z \right\}. \quad (5-17)$$

A solução da equação de Helmholtz em coordenadas cilíndricas obtida por separação de variáveis é igual a [59, p. 113]:

$$v(\rho, \theta, z) = f(\kappa_\rho \rho) \times \left(C_v e^{im\theta} + D_v e^{-im\theta} \right) \times \left(E_v e^{-i\kappa_z z} + F_v e^{+i\kappa_z z} \right), \quad (5-18)$$

com

$$\kappa^2 = \kappa_\rho^2 + \kappa_z^2. \quad (5-19)$$

Além disso, m é uma constante inteira não-negativa e f é uma função da coordenada ρ que ficará clara nos próximos parágrafos. Por fim, as demais constantes são escalares complexos.

Portanto, de 5-18 deduz-se que $v_{[zz]} = -\kappa_z^2 v$, e usando 5-19 conclui-se que 5-17 é igual a

$$\mathbf{E}_{\text{tm}} = \kappa^{-1} \left\{ v_{[z\rho]} \mathbf{a}_\rho + \rho^{-1} v_{[z\theta]} \mathbf{a}_\theta + \kappa_\rho^2 v \mathbf{a}_z \right\}. \quad (5-20)$$

Analogamente, os campos magnéticos em 5-15 ficam

$$\mathbf{H}_{\text{tm}} = -\kappa \tau^{-1} \left(v_{[\theta]} \mathbf{a}_\rho - v_{[\rho]} \mathbf{a}_\theta \right), \quad (5-21)$$

$$\mathbf{H}_{\text{te}} = -\tau^{-1} \left\{ u_{[z\rho]} \mathbf{a}_\rho + \rho^{-1} u_{[z\theta]} \mathbf{a}_\theta + \kappa_\rho^2 u \mathbf{a}_z \right\}, \quad (5-22)$$

com u definido de maneira semelhante a 5-18.

O tubo de produção comporta-se como um guia de ondas circular não homogêneo na direção radial, devido às diferentes camadas de materiais dispostos de maneira circular e concêntrica em torno do eixo de simetria do tubo. Por isso, os próximos parágrafos mostram uma formulação geral para calcular os modos de propagação em um guia de ondas cilíndrico radialmente fragmentado em $l + 1$ camadas. Em particular, a camada de índice $j = 0$ é a camada central e a camada de índice $j = l$ é a camada mais externa, de tamanho infinito, como mostra a figura 12. Seja ρ_j o raio menor da camada j e considere que cada camada j , $0 \leq j \leq l$ satisfaz $\rho_j \leq \rho < \rho_{j+1}$. Sendo assim, $\rho_0 = 0$, $\rho_{l+1} = \infty$ e os demais valores de ρ_j devem ser conhecidos de antemão.

Na interface de separação entre duas camadas adjacentes j e $j + 1$, $0 \leq j < l$, os campos elétrico e magnético devem ser contínuos nas componentes tangenciais \mathbf{a}_θ e \mathbf{a}_z [60, 61]. Logo, a partir das equações 5-12, 5-16, 5-20 e 5-22,

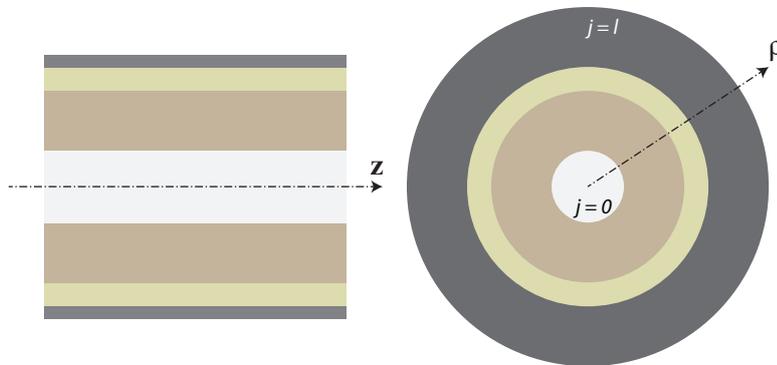


Figura 12: Representação de um guia de ondas dividido em 4 camadas, sendo a última semi-infinita.

as seguintes relações devem ser satisfeitas para cada $\rho = \rho_j$, $0 \leq j < l$:

$$-u_{[\rho]}^j + \kappa_j^{-1} \rho_{j+1}^{-1} v_{[z\theta]}^j = -u_{[\rho]}^{j+1} + \kappa_{j+1}^{-1} \rho_{j+1}^{-1} v_{[z\theta]}^{j+1} \quad (5-23)$$

$$(\kappa_\rho^j)^2 \kappa_j^{-1} v^j = (\kappa_\rho^{j+1})^2 \kappa_{j+1}^{-1} v^{j+1} \quad (5-24)$$

$$\kappa_j \tau_j^{-1} v_{[\rho]}^j - \kappa_j \tau_j^{-1} \rho_{j+1}^{-1} u_{[z\theta]}^j = \kappa_{j+1} \tau_{j+1}^{-1} v_{[\rho]}^{j+1} - \kappa_{j+1} \tau_{j+1}^{-1} \rho_{j+1}^{-1} u_{[z\theta]}^{j+1} \quad (5-25)$$

$$-\tau_j^{-1} (\kappa_\rho^j)^2 u_j = -\tau_{j+1}^{-1} (\kappa_\rho^{j+1})^2 u_{j+1}. \quad (5-26)$$

Nas relações acima, as constantes q_1 e q_2 foram embutidas nas funções u e v . Além disso, a notação com índice j sobrescrito indica o objeto (função ou constante) referente à camada de índice j .

5.2.3

Dedução do Problema de Autovalores

Motivado pela equação 5-18, podemos buscar soluções na forma

$$\begin{aligned} v &= \hat{v}(\rho) e^{i(m\theta - \kappa_z z)}, \\ u &= \hat{u}(\rho) e^{i(m\theta - \kappa_z z)}, \\ z &\geq 0, \\ \kappa_z &\in \mathbb{C} \\ m &= 0, 1, 2, \dots \\ \text{Re } \kappa_z &> 0, \quad \text{Im } \kappa_z \leq 0. \end{aligned} \quad (5-27)$$

Onde \hat{v} e \hat{u} são funções de uma dimensão. Assim, as equações 5-10 e 5-11 podem ser reescritas como

$$\begin{aligned} \rho^{-1} (\rho \hat{u}_{[\rho]})_{[\rho]} - m^2 \rho^{-2} \hat{u} - \kappa_z^2 \hat{u} + \kappa^2(\rho) \hat{u} &= 0 \\ \rho^{-1} (\rho \hat{v}_{[\rho]})_{[\rho]} - m^2 \rho^{-2} \hat{v} - \kappa_z^2 \hat{v} + \kappa^2(\rho) \hat{v} &= 0, \end{aligned} \quad (5-28)$$

onde $\kappa^2(\rho)$ foi usado para enfatizar que a constante κ^2 varia na direção radial por causa dos diversos meios de propagação (óleo, gás, incrustação, metal, rocha). Em notação de operadores, conforme mostrado no capítulo 3, o

problema pode ser escrito como o problema de autovalores

$$\begin{aligned}
 A\mathbf{h} &= \lambda\mathbf{h}, \\
 \mathbf{h} &= [\hat{u} \ \hat{v}]^T, \\
 A\mathbf{h} &= \begin{bmatrix} \rho^{-1}(\rho\hat{u}_{[\rho]})_{[\rho]} - m^2\rho^{-2}\hat{u} + \kappa^2(\rho)\hat{u} \\ \rho^{-1}(\rho\hat{v}_{[\rho]})_{[\rho]} - m^2\rho^{-2}\hat{v} + \kappa^2(\rho)\hat{v} \end{bmatrix}, \\
 \lambda &= \kappa_z^2 \\
 u, v &\in L^2([0; +\infty[).
 \end{aligned} \tag{5-29}$$

É conveniente dividir o problema 5-29 em subproblemas referentes à cada meio de propagação, conectados entre si pelas condições de contorno. Assim, a versão final do problema de autovalores é

$$\begin{aligned}
 \tilde{A}\mathbf{H} &= \lambda\mathbf{H}, \\
 \tilde{A}\mathbf{H} &= [A\mathbf{h}_0; \ A\mathbf{h}_1; \ \dots; \ A\mathbf{h}_l] \\
 \mathbf{H} &= [\mathbf{h}_0; \ \dots; \ \mathbf{h}_l] \\
 \mathbf{h}_j &= [u_j; \ v_j], \\
 \hat{u}_j(\rho) &= \begin{cases} \hat{u}(\rho), & \rho_j \leq \rho < \rho_{j+1} \\ 0, & \text{caso contrário} \end{cases} \\
 \hat{v}_j(\rho) &= \begin{cases} \hat{v}(\rho), & \rho_j \leq \rho < \rho_{j+1} \\ 0, & \text{caso contrário} \end{cases} \\
 \lambda &= \kappa_z^2 \\
 \hat{u}, \hat{v} &\in L^2([0; +\infty[) \\
 \mathbf{H} &\in L^2([0; +\infty[)^{2(l+1)},
 \end{aligned} \tag{5-30}$$

onde o “;” indica concatenação vertical para formar uma matriz coluna.

Note que a função vetorial H tem contradomínio no $\mathbb{C}^{2(l+1)}$ onde $l + 1$ é o número de meios de propagação. Neste caso, o produto interno usado é

$$\begin{aligned}
 \langle \mathbf{f}; \mathbf{g} \rangle &= \int_0^{+\infty} \mathbf{f} \cdot \bar{\mathbf{g}} \rho \, d\rho, \\
 \mathbf{f}, \mathbf{g} &\in L^2([0; +\infty[)^{2(l+1)}
 \end{aligned} \tag{5-31}$$

e “.” indica o produto escalar do $\mathbb{R}^{2(l+1)}$, sem conjugação.

Além da definição apresentada na equação 5-30, a função vetorial \mathbf{H} também deve satisfazer ao conjunto de condições de contorno especificados nas equações 5-23, 5-24, 5-25 e 5-26. No total, são $4l$ condições de contorno,

que podem ser expressas em notação de operadores como

$$\begin{aligned} A_0 \mathbf{H} + k_z A_1 \mathbf{H} &= 0 \\ \mathbf{H} &\in L^2([0; +\infty])^{2(l+1)} \\ A_0 \mathbf{H}, A_1 \mathbf{H} &\in L^2([0; +\infty])^{4l+2(l+1)}. \end{aligned} \quad (5-32)$$

Note que o contradomínio dos operadores A_0 e A_1 é um espaço de funções vetoriais com $4l+2(l+1)$ dimensões. Essa estratégia foi adotada para incluir as $2(l+1)$ equações em 5-30 e as $4l$ equações referentes às condições de contorno.

Ao juntar 5-32 com o problema de autovalores 5-30, temos o problema de autovalores quadrático

$$\begin{aligned} A_0 \mathbf{H} + k_z A_1 \mathbf{H} + k_z^2 A_2 \mathbf{H} &= 0 \\ \mathbf{H} &\in L^2([0; +\infty])^{2(l+1)}, \end{aligned} \quad (5-33)$$

onde A_0, A_1 e A_2 são operadores com domínio igual a $L^2([0; +\infty])^{2(l+1)}$ e contradomínio $L^2([0; +\infty])^{4l+2(l+1)}$. Neste tipo de problema, k_z é o autovalor a ser encontrado e \mathbf{H} é a autofunção. Ao aplicar o produto interno em ambos os lados da eq. 5-33, temos

$$\begin{aligned} \langle A_0 \mathbf{H}; A_0 \mathbf{H} \rangle' + k_z \langle A_1 \mathbf{H}; A_0 \mathbf{H} \rangle' + k_z^2 \langle A_2 \mathbf{H}; A_0 \mathbf{H} \rangle' &= 0 \\ \mathbf{H} &\in L^2([0; +\infty])^{2(l+1)} \\ k_z &\in \mathbb{C}, \end{aligned} \quad (5-34)$$

onde $\langle \cdot; \cdot \rangle'$ foi usado para indicar produto interno entre funções de $4l+2(l+1)$ dimensões. A equação 5-34 é uma equação do segundo grau na variável k_z e, portanto, dada a autofunção \mathbf{H} , calcula-se k_z usando 5-34.

Portanto, temos a formulação matemática do problema de autovalores a ser resolvido. Os autopares de 5-34 são os modos de propagação da onda eletromagnética no interior do poço de petróleo. Além disso, os operadores A_0, A_1 e A_2 não são normais, logo temos o primeiro estudo de caso em que o modelo proposto no capítulo 3 é usado para resolver este tipo de problema. Nas próximas seções, será apresentado um modelo semi-analítico para resolver este problema de autovalores, visto que uma solução totalmente analítica não é possível. Este modelo semi-analítico será então comparado ao modelo proposto para um estudo de caso típico de um poço de petróleo e conclusões acerca do modelo proposto serão apresentadas.

5.3

Modelo semi-analítico de solução

Considere novamente 5-18. Na camada central, que tem $j = 0$, a solução que é não-singular em todos os pontos $0 \leq \rho \leq \rho_1$ tem:

$$f(\kappa_\rho \rho) = J_m(\kappa_\rho \rho), \quad (5-35)$$

onde J_m é a função de Bessel de primeira espécie de ordem m . Este trabalho também usa a notação Y_m para a função de Bessel de segunda espécie.

A solução na camada semi-infinita $j = l$, válida para $\rho \geq \rho_l$, deve representar uma onda que se propaga na direção $+\mathbf{a}_\rho$. Por isso, a função de Hankel de segunda espécie $H_m^{(2)}$, definida para todo complexo x como $H_m^{(2)}(x) = J_m(x) - iY_m(x)$, é a mais adequada [59, p. 115]. Logo, para $\rho \geq \rho_l$,

$$f(\kappa_\rho \rho) = H_m^{(2)}(\kappa_\rho \rho), \quad (5-36)$$

Para as demais camadas j , $0 < j < l$, a solução f_ρ deve representar uma onda estacionária na direção \mathbf{a}_ρ . Por isso, este modelo usa

$$f(\kappa_\rho \rho) = AJ_m(\kappa_\rho \rho) + BY_m(\kappa_\rho \rho), \quad (5-37)$$

para todo ρ tal que $\rho_1 \leq \rho < \rho_l$, onde A e B são constantes complexas.

Utilizando 5-18 para a camada j com $\rho_j \leq \rho < \rho_{j+1}$, quando $\rho \rightarrow \rho_{j+1}$, temos:

$$\begin{aligned} v_j &= f_v(\kappa_\rho \rho_{j+1}) \times e^{i(m\theta - \kappa_z z)} \\ u_j &= f_u(\kappa_\rho \rho_{j+1}) \times e^{i(m\theta - \kappa_z z)}. \end{aligned} \quad (5-38)$$

Aplicando essas expressões às equações 5-23, 5-24, 5-25 e 5-26 para todas as interfaces $j \rightarrow j + 1$, $j = 1, 2, \dots, l$, obtém-se um sistema linear homogêneo:

$$M\mathbf{w} = 0 \quad (5-39)$$

onde M é uma matriz quadrada de ordem $4l$ que depende de κ_z e \mathbf{w} é um vetor formado pelos coeficientes de v_j e u_j . Portanto, para que o sistema 5-39 tenha soluções não nulas, o determinante de sua matriz deve ser igual a 0, isto é:

$$D = \det M = 0. \quad (5-40)$$

Portanto, o problema se reduz a encontrar as raízes da equação univariada $D(\kappa_z) = 0$ no interior de um retângulo do plano complexo.

Existem algumas abordagens na literatura para encontrar todas as raízes de equações não lineares no plano complexo. Em [62], os autores apresentaram uma metodologia baseada em dois testes para verificar a existência de raízes. O domínio original é um retângulo especificado pelo usuário e, caso existam raízes dentro deste domínio, os autores o particionam em 4 subdomínios e continuam a análise recursivamente. Quando um subdomínio atinge um tamanho mínimo, os autores aplicam o método de Newton para encontrar a raiz. Tal metodologia funciona apenas para encontrar raízes isoladas, logo não é adequada quando raízes múltiplas são possíveis dentro do retângulo de busca original.

Em contraposição à [62], os autores de [63] desenvolveram uma heurística com rigor matemático que garante que todas as raízes da equação $D(\kappa_z) = 0$, múltiplas ou isoladas, sejam de fato encontradas de maneira eficiente. Os autores assumem que o domínio é um retângulo R e que a função D é analítica em todos os pontos de R . Além disso, não devem existir raízes de D no contorno de R .

O método de [63] baseia-se numa versão do princípio do argumento para funções analíticas. Basicamente, se D é uma função analítica no retângulo R , então o número de raízes no interior de R é igual a:

$$N_r(D, R) = \frac{1}{2\pi i} \int_{\partial R} \frac{D'(x)}{D(x)} dx, \quad (5-41)$$

onde ∂R é o contorno de R e a integral acima é uma integral de linha. Portanto, dado $N_r(D, R)$, três cenários são possíveis: (1) $N_r(D, R) = 0$, logo R não têm raízes; (2) $N_r(D, R) = 1$, logo existe uma única raiz no interior de R ; e (3) $N_r(D, R) > 1$, ou seja, existem várias raízes em R , possivelmente com multiplicidade superior a 1.

No caso de uma única raiz, a heurística aplica o método de Newton para achá-la. Caso o método de Newton falhe em encontrar a raiz, ele procede particionando R em duas partes e todo o procedimento repete-se para cada parte até que a raiz seja encontrada ou algum critério limite seja alcançado.

Por outro lado, caso $N_r(D, R) > 1$, temos duas possibilidades: o interior do retângulo tem $N_r(D, R)$ raízes idênticas, ou R apresenta pelo menos 1 par de raízes distintas. No primeiro caso, particionar o domínio não reduz o número de raízes, logo basta aplicar uma abordagem semelhante ao caso $N_r(D, R) = 1$. Em contraposição, no caso com pelo menos 1 par de raízes distintas, é necessário continuar particionando para isolar estas raízes umas das outras. Para distinguir entre os dois casos, os autores de [63] optam por aplicar o teste da equação 5-41 na derivada de ordem $N_r(D, R) - 1$ da função D e, caso o resultado seja nulo, então necessariamente a derivada de ordem

$N_r(D, R) - 1$ nunca será nula em R . Neste caso, a função D tem no máximo $N_r(D, R) - 1$ raízes idênticas no interior do conjunto R ou, equivalentemente, existe pelo menos 1 par de raízes distintas.

Por fim, os autores ainda sugerem utilizar aritmética intervalar para calcular a integral 5-41 com o menor número de iterações possível mas que ainda garanta que o resultado esteja correto. Como o resultado de 5-41 é sempre um inteiro não negativo, basta garantir que a integral seja calculada de modo que o erro da parte real seja menor do que $1/2$ para evitar que seu resultado seja confundido com o valor de algum inteiro adjacente. Embora essa abordagem funcione bem, ela não é diretamente aplicável à equação 5-40 uma vez que a aritmética intervalar complexa aplicada às funções de Bessel não é trivial.

Um outro ponto fraco da heurística apresentada em [63] é que, no caso de existirem raízes no contorno de R , a integral 5-41 não converge para um número inteiro. Embora o retângulo inicial seja pré-fornecido pelo usuário, os sub-retângulos calculados durante a execução do algoritmo podem sim ter raízes numericamente muito próximas de seus contornos, o que inviabilizaria o princípio do argumento como um teste da presença de raízes.

Sendo assim, este trabalho utiliza uma versão modificada do algoritmo de [63], que não usa a aritmética intervalar e que trata com cuidado o problema da existência de raízes no contorno das regiões de busca. Para isto, sempre que a integral 5-41 não convergir para um inteiro em um número de passos pequeno, o domínio R é ampliado por um fator de escala ligeiramente maior do que 1, e repete-se o processo.

O algoritmo 4 resume o procedimento para encontrar as raízes. Ele utiliza uma fila *FIFO* (*first in first out*) para armazenar os retângulos a serem utilizados na busca das raízes. Para cada iteração, o algoritmo busca o próximo domínio, $prox_ret$, da fila Q e testa se ele é muito pequeno. Se sim, procede-se com o método de Newton para achar alguma raiz no interior deste domínio e o algoritmo continua para a próxima iteração, sem particionar este domínio. Caso $prox_ret$ não seja de tamanho inferior ao limiar, o algoritmo procede aplicando o princípio do argumento (equação 5-41) à função D para calcular o número de raízes no interior de $prox_ret$. Se existir uma única raiz, utiliza-se o método de Newton para busca-la e coloca-la na lista L . Caso o método de Newton não encontre a raiz, a função auxiliar $BUSCAR_NEWTON(prox_ret, L, Q)$ particiona $prox_ret$ em duas metades, adicionando-as à Q . Finalmente, se existir mais do que 1 raiz, o algoritmo verifica a possibilidade de tratar-se de uma única raiz de multiplicidade W ou não. No primeiro caso, $Y = 1$, busca-se a raiz usando novamente a função auxiliar $BUSCAR_NEWTON(prox_ret, L, Q)$ e no

Algoritmo 4 Raízes de uma função complexa

```

1: procedure ACHAR_RAÍZES( $R, d_{min}$ ) retorna  $L$ .
2:    $L \leftarrow \{\}, Q \leftarrow \{R\}$ .
3:   repeat
4:      $prox\_ret \leftarrow$  próximo de  $Q$ .
5:     if  $DIAM(prox\_ret) < d_{min}$  then
6:        $BUSCAR\_NEWTON(prox\_ret, L, nil)$ 
7:       continue para a próxima iteração.
8:     Calcular  $W = N_r(D, prox\_ret)$ .
9:     if  $W = 0$  then
10:      continue para a próxima iteração.
11:    else if  $W = 1$  then
12:       $BUSCAR\_NEWTON(prox\_ret, L, Q)$ 
13:    else
14:      Calcular  $Y = N_r(D^{(W-1)}, prox\_ret)$ .
15:      if  $Y = 0$  then
16:        Enfileire em  $Q$  duas metades de  $prox\_ret$ .
17:      else
18:         $BUSCAR\_NEWTON(prox\_ret, L, Q)$ 
19:    until  $Q$  esteja vazia.

```

segundo caso o algoritmo subdivide R em duas metades, enfileira-as em Q e procede para a próxima iteração.

Para deixar mais claro, o algoritmo 5 mostra o funcionamento da função auxiliar $BUSCAR_NEWTON(prox_ret, L, Q)$.

Algoritmo 5 Função auxiliar para achar raiz

```

1: procedure  $BUSCAR\_NEWTON(R, L, Q)$ 
2:    $r \leftarrow$  raiz de  $D$  no interior de  $R$  usando método de Newton.
3:   if  $r \neq nil$  then
4:     Adiciona a raiz  $r$  à lista  $L$ .
5:   else
6:     if  $Q \neq nil$  then
7:       Particione  $R$  em duas metades e adicione-as à  $Q$ .

```

Resta, portanto, explicar em detalhes como calcular a integral da equação 5-41. Primeiramente, reescreve-se a integral como uma soma de quatro integrais nos 4 lados do retângulo:

$$N_r(D, R) = \frac{1}{2\pi i} (I_1 + I_2 + I_3 + I_4), \quad (5-42)$$

onde

$$I_j = \int_0^1 \frac{D'(x_j(t))}{D(x_j(t))} \frac{dx_j}{dt} dt, \quad (5-43)$$

para cada lado do retângulo indexado por $j = 1, 2, 3, 4$. Para um determinado lado j com vértices inicial $x_j^{(-)}$ e final $x_j^{(+)}$, tem-se $x_j(t) = x_j^{(-)} + (x_j^{(+)} - x_j^{(-)})t$ com $0 \leq t \leq 1$.

A fim de calcular a integral 5-43, aplica-se integração numérica usando a regra de Simpson:

$$\int_a^b g(t)dt = \frac{b-a}{6} \left(g(a) + g(b) + 4g\left(\frac{a+b}{2}\right) \right) + \text{erro}$$

$$\text{erro} = -\frac{1}{90} \left(\frac{b-a}{2}\right)^5 g^{(4)}(\chi), \quad (5-44)$$

na qual χ é um número entre a e b . A última parcela representa o erro de aproximação, cuja parte real deve ter valor absoluto menor do que $\pi/4$ para garantir que o erro acumulado na aproximação de 5-42 tenha parte real menor do que $1/2$. O processo de integração é recursivo, particionando o intervalo de integração até que garantidamente o erro de integração do intervalo $[0; 1]$ seja menor do que $1/2$. Como o erro decai com a quinta potência de $b - a$, rapidamente converge-se para a solução.

5.4

Modelagem eletromagnética dos materiais

O problema aqui descrito tem ondas propagando-se em óleo, gás natural, carbonato de cálcio, aço e rocha. Por isso, esta seção dedica-se a resumir os modelos usados para as relações constituintes e as propriedades de cada meio de propagação.

Todos os referidos meios, exceto o aço, apresentam permeabilidade magnética semelhante à permeabilidade do vácuo, logo utiliza-se $\mu = \mu_0$. Para o caso particular do aço, por ser um metal, utiliza-se uma permeabilidade magnética bem maior, da ordem de $500\mu_0$. É assumido neste trabalho que não há efeitos dispersivos na permeabilidade magnética em nenhum dos meios.

Em [64, p. 323], os autores relatam que o CaCO_3 tem características anisotrópicas para frequências da ordem de 94 GHz, mas para frequências da ordem de centenas de MHz esta anisotropia é bem menos notável. Por isso, assume-se que o CaCO_3 é um material isotrópico com permissividade relativa igual a 8,5. Além disso, assume-se que a condutividade do CaCO_3 é nula.

Em [65], os autores relatam que os gases metano, etano, propano e butano, que são os principais componentes do gás natural, apresentam constante dielétrica relativa muito próxima de 1, a não ser que a densidade destes gases seja muito alta. Por isso, este trabalho assume que o gás natural tem permissividade elétrica igual à do vácuo. Assim como no CaCO_3 assume-se

Tabela 18: Constantes do meio de propagação “óleo”

Amostra	ϵ_s	ϵ_∞	ψ (ns)	a	σ (nS/m)
D	2,341	2,197	14,65	0,500	2,51
F	2,277	2,167	2,34	0,459	9,34
G	2,325	2,194	4,52	0,494	9,43

Tabela 19: Constantes dos meios de propagação

Meio	ϵ	μ	σ
óleo cru	tabela 18	μ_0	tabela 18
aço	ϵ_0	$500\mu_0$	$10^6 S/m$
CaCO ₃	$8,5\epsilon_0$	μ_0	0
gás natural	ϵ_0	μ_0	0
rocha	ϵ_0	μ_0	$3,5 \times 10^{-2} S/m$
fluido de completação	$3\epsilon_0$	μ_0	$1,0 \times 10^{-9} S/m$

condutividade nula para o gás natural.

Em [66], os autores apresentam um modelo de Cole-Cole para diferentes misturas de óleo e água, obtidas diretamente de um reservatório real, sem nenhum tipo de separação. Eles escrevem a constante dielétrica da mistura como

$$\epsilon = \epsilon_0 \left[\epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + (i\omega\psi)^{1-a}} \right], \quad (5-45)$$

onde ϵ_s e ϵ_∞ são as constantes dielétricas relativas em $\omega = 0$ e $\omega = +\infty$, ψ é o tempo de relaxação do material e a é um fator chamado fator de distribuição. O autor apresenta os valores destas constantes, juntamente com a condutividade elétrica, para três amostras distintas de óleo misturado com água. A tabela 18 reproduz os valores publicados em [66].

Finalmente, no caso do aço utilizado no tubo de produção e nos revestimentos, utiliza-se uma condutividade de $10^6 S/m$. A tabela 19 resume a discussão desta seção.

5.5 Resultados

Esta seção apresenta uma análise dos resultados obtidos pelo método semi-analítico e pelo método proposto neste trabalho. Para isso, assume-se uma configuração típica de poço, com raio variando de 5 a 10 cm, 10% a 25% de incrustação de CaCO₃, e uma *GOR* fixa em 2. Além disso, assume-se que o tubo de produção tem espessura de 5 mm; a região anular, formada por fluido de completação, tem espessura de 3 cm, e o revestimento externo do poço tem espessura de 10 mm.

O modelo semi-analítico é usado apenas a título de comparação. Embora ele tenha como ponto forte a alta performance devido ao número reduzido de computações numéricas, ele não é paramétrico. Por esta razão, ele não é adequado para analisar o efeito de várias combinações de incrustações, relação gás óleo e raio do poço, pois demandaria um número grande de simulações. Por outro lado, o modelo paramétrico proposto neste trabalho se adéqua a este tipo de aplicação, pois é capaz de entregar como resposta autofunções parametrizadas que aproximam os autovalores para qualquer valor prático de raio de poço, relação gás-óleo e percentual de incrustação.

Devido ao decaimento exponencial da função de Hankel, $H_m^{(2)}(x)$, com o aumento do valor absoluto de x , é muito difícil executar numericamente o modelo semi-analítico. Isso ocorre porque o computador, com sua precisão finita, trunca para 0 o valor de $H_m^{(2)}(\kappa_\rho^l \rho_l)$ durante a montagem do sistema de equações 5-39.

Assim, ao invés de trabalhar com $H_m^{(2)}(x)$, optou-se por trabalhar com a função $f_l(\rho) = H_m^{(2)}(\kappa_\rho^l \rho) / H_m^{(2)}(\kappa_\rho^l \rho_l)$. No contorno $\rho = \rho_l$, $f(\rho) = 1$ o que elimina os problemas de aproximação. Além disso, para calcular sua derivada avaliada no contorno $f_{l[\rho]}(\rho_l)$, utiliza-se uma série de Laurent adequada para valores grandes do argumento, expandida em torno de infinito. Alguns exemplos incluem

$$\begin{aligned} \frac{H_0^{(2)}_{[\rho]}(\kappa_\rho^l \rho_l)}{H_0^{(2)}(\kappa_\rho^l \rho_l)} &= -i - \frac{1}{2x} - \frac{i}{8x^2} + \frac{1}{8x^3} + \frac{25i}{128x^4} - \frac{13}{32x^5} \\ \frac{H_1^{(2)}_{[\rho]}(\kappa_\rho^l \rho_l)}{H_1^{(2)}(\kappa_\rho^l \rho_l)} &= -i - \frac{1}{2x} + \frac{3i}{8x^2} - \frac{3}{8x^3} - \frac{63i}{128x^4} + \frac{27}{32x^5} \\ x &= \kappa_\rho^l \rho_l. \end{aligned} \tag{5-46}$$

Essas aproximações apresentam erro menor que 10^{-10} para valores do argumento com módulo maior do que 10, o que foi confirmado utilizando computação simbólica via software Maple. Na prática, $\kappa_\rho^l \rho_l$ é da ordem de 10^5 devido ao alto valor de condutividade do metal do tubo de produção.

Para o modelo proposto, utilizou-se as configurações apresentadas na tabela 20. O critério de parada foi de 1000 gerações. Além disso, usou-se o critério de remoção de autofunções conhecidas apresentado na seção 3.2.5.

5.5.1

Tabela 20: Modelo proposto para achar os modos do poço de petróleo.

População	50 indivíduos
Gerações	1000 gerações
Núm. Árvores (k_a)	2 por função escalar
Taxa de mutação	fixa em 10%
Tipos de mutação	6 tipos descritos no cap. 3
Tipo de recombinação	sub-árvore
Funções	$\{e^{-t}, \text{Bessel de ordem } 0 \text{ a } 2, \text{sen}(t), \text{cos}(t), *, +\}$
Terminais	$\{\rho, \kappa, 0, 1, -1, i\}$
Otimizador Local	<i>BFGS Quasi-Newton</i> [34, 35]
Avaliação	alg. 1 + derivadas e simplificação simbólicas
Profundidade Máxima	5 níveis por árvore

Tabela 21: Valor de κ_z para os primeiros modos com $\kappa_z \in [0, 1 - 150i; 100 - 10^{-4}i]$, frequência 0,8 GHz, óleo tipo “D”, por ordem crescente de atenuação.

	m	semi-analítico (m^{-1})	proposto (m^{-1})	mdf (m^{-1})
1	1	$11,8485 - 0,0114i$	$11,9179 - 0,0120i$	$11,8877 - 0,0118i$
2	2	$11,1971 - 0,0316i$	$11,1903 - 0,0299i$	$11,2001 - 0,0377i$
3	3	$11,1904 - 0,0336i$	$11,1939 - 0,0381i$	$11,3014 - 0,0401i$
4	0	$24,9184 - 0,0568i$	$25,0103 - 0,0577i$	$25,0101 - 0,0481i$
5	1	$47,0717 - 0,1233i$	$47,2044 - 0,1200i$	$50,6157 - 0,1001i$
6	2	$65,7582 - 0,2628i$	$66,0511 - 0,2566i$	$65,2204 - 0,2447i$
7	3	$83,6606 - 0,5529i$	$83,6022 - 0,5178i$	$85,0200 - 0,5491i$
8	0	$0,1128 - 72,8348i$	$0,1293 - 70,9356i$	$0,1027 - 72,6611i$
9	0	$0,1826 - 138,3167i$	$0,2254 - 133,2346i$	$0,1770 - 151,2334i$

Experimento não-paramétrico

O primeiro experimento é não-paramétrico, assumindo raio do poço igual a 5 cm e 10% de incrustação de CaCO_3 . Ele foi executado para duas faixas de frequência: 0,8 GHz e 1 kHz. Neste experimento, espera-se que as técnicas apresentem mais dificuldade para resolver o problema em baixa frequência, pois, em domínios infinitos, as integrais presentes nos produtos internos demandam mais pontos para convergirem em baixas frequências do que em altas frequências. Além disso, em métodos baseados em discretização, é necessário truncar a malha de pontos, o que gera perda de acurácia. Por esta razão, este experimento também considera o Método das Diferenças Finitas a título de comparação, de modo a mostrar as vantagens do método proposto sobre métodos baseados em discretização em domínios infinitos.

A tabela 21 mostra os primeiros modos calculados com o método semi analítico e com o modelo proposto para uma frequência $\omega = 2\pi \cdot 0,8 \text{ GHz}$. Neste experimento, buscou-se valores de k_z no retângulo complexo $[0, 1 - 150i; 100 - 10^{-4}i]$.

Tabela 22: Valor de κ_z para os primeiros modos com $\kappa_z \in [0, 1 - 150i; 100 - 10^{-4}i]$, frequência 0,8 GHz, óleo tipo “F”, por ordem crescente de atenuação.

	m	semi-analítico (m^{-1})	proposto (m^{-1})	mdf (m^{-1})
1	1	10,7701 - 0,0187i	12,0001 - 0,0191i	12,0045 - 0,0188i
2	2	12,2720 - 0,0371i	12,4545 - 0,0370i	13,4504 - 0,0391i
3	3	14,4200 - 0,0440i	14,6790 - 0,0488i	14,3025 - 0,0404i
4	0	29,4724 - 0,0667i	29,5501 - 0,0671i	30,5015 - 0,0690i
5	1	51,9681 - 0,1488i	51,9700 - 0,1491i	50,6157 - 0,1477i
6	2	77,6640 - 0,2908i	76,9991 - 0,2901i	78,0186 - 0,3016i
7	3	101,2221 - 0,5711i	102,0117 - 0,5781i	102,7530 - 0,5810i
8	0	0,1001 - 74,7800i	0,1293 - 70,9356i	0,1152 - 77,1118i
9	0	0,1912 - 175,5378i	0,2170 - 178,7737i	0,2122 - 171,2452i

Tabela 23: Valor de κ_z para os primeiros modos com $\kappa_z \in [0, 1 - 150i; 100 - 10^{-4}i]$, frequência 0,8 GHz, óleo tipo “G”, por ordem crescente de atenuação.

	m	semi-analítico (m^{-1})	proposto (m^{-1})	mdf (m^{-1})
1	1	12,1004 - 0,0191i	12,0022 - 0,0192i	12,0109 - 0,0189i
2	2	12,2829 - 0,0377i	12,3607 - 0,0377i	12,9973 - 0,0383i
3	3	14,4512 - 0,0497i	14,4880 - 0,0495i	14,4776 - 0,0492i
4	0	29,7050 - 0,0699i	29,77561 - 0,0700i	30,1102 - 0,0710i
5	1	52,0074 - 0,1501i	52,0200 - 0,1522i	51,2330 - 0,1510i
6	2	79,9880 - 0,3033i	79,5647 - 0,3101i	79,1042 - 0,3110i
7	3	101,4920 - 0,5872i	101,9400 - 0,5819i	102,4705 - 0,5855i
8	0	0,1033 - 79,5894i	0,1075 - 79,7822i	0,1134 - 79,8540i
9	0	0,2004 - 179,1247i	0,2780 - 179,4577i	0,2441 - 178,9389i

Os resultados da tabela 21 mostram na primeira linha o modo de propagação com atenuação mínima para a configuração usada. Em particular, este experimento usou óleo do tipo “D” (tabela 18) e frequência de 0,8 GHz.

Observando a tabela 21, conclui-se que o modelo proposto conseguiu encontrar resultados muito semelhantes ao modelo semi-analítico. Do ponto de vista de resultados, a atenuação menor (parte imaginária de κ_z) ocorre no modo identificado como 1 e seu coeficiente de atenuação é próximo de $-0.0114m^{-1}$ de acordo com os modelos. Isso implica em uma atenuação de amplitude do campo elétrico da ordem de $10\log(e^{-0.0114}) = 0.0495$ dB/m. Boa parte dessa atenuação se deve às perdas no óleo, mas as perdas no metal por efeito Joule também são consideráveis.

Os modos 8 e 9 apresentados na tabela 21 apresentam atenuação altíssima. Isso ocorre porque a frequência de propagação de 0.8 GHz está abaixo da frequência mínima de ativação destes modos, também conhecida como frequência de corte. Por esta razão, tais modos não se propagam, embora sejam úteis para fazer casamento de ondas incidentes, refletidas e refratadas na interface de separação de dois meios.

Tabela 24: Erro de cada um dos métodos para obtenção dos modos apresentados na tabela 21.

Modo	m	semi-analítico	proposto	mdf
1	1	1.874×10^{-4}	1.702×10^{-4}	3.105×10^{-4}
2	2	1.260×10^{-4}	1.673×10^{-4}	3.948×10^{-4}
3	3	1.368×10^{-4}	1.237×10^{-4}	2.977×10^{-4}
4	0	1.510×10^{-4}	1.647×10^{-4}	3.684×10^{-4}
5	1	2.102×10^{-4}	1.869×10^{-4}	4.005×10^{-4}
6	2	2.017×10^{-4}	1.988×10^{-4}	3.978×10^{-4}
7	3	2.337×10^{-4}	2.191×10^{-4}	6.804×10^{-4}
8	0	2.294×10^{-4}	2.101×10^{-4}	6.748×10^{-4}
9	0	2.933×10^{-4}	2.001×10^{-4}	10.221×10^{-4}

Outro ponto importante é que o modo 4 tem $\kappa_\rho = 0$ no óleo. Este modo está relacionado ao modo transversal eletromagnético em guias de onda coaxiais ideais e uma característica importante é que ele não tem frequência de corte, isto é, este modo se propaga em qualquer frequência. Este modo surge fazendo $\kappa_\rho = 0$ na solução da equação de Helmholtz para o óleo, o que leva à equação diferencial de Euler na direção radial, ao invés da equação diferencial de Bessel.

Em termos de acurácia, a tabela 24 apresenta o erro, definido como $\epsilon = \frac{\|A_0 \mathbf{H} + \kappa_z A_1 \mathbf{H} + \kappa_z^2 A_2 \mathbf{H}\|}{\|\mathbf{H}\|}$, para cada célula da tabela 21. Os resultados mostram que o Método das Diferenças Finitas obteve o pior desempenho, devido à necessidade de truncar a malha de pontos para ρ suficientemente grande. Neste caso, optou-se por truncar para $\rho > 2\lambda_0$, onde λ_0 é o comprimento de onda da luz no vácuo na frequência do experimento. Neste caso, $\lambda_0 = 3 \times 10^8 / 0,8 \times 10^9 = 0,375\text{m}$, donde conclui-se que a malha foi truncada para $\rho > 0,75\text{m}$.

Além disso, os resultados da tabela 24 também mostram uma superioridade da técnica proposta, por ter apresentado geralmente menos erro, sobretudo para modos mais elevados. Adicionalmente, note que a técnica proposta, baseada em PG, é a única capaz de entregar soluções analíticas, o que por si só já seria uma vantagem sobre as técnicas usadas como comparação.

Também foram realizados experimentos envolvendo os óleos dos tipos “F” e “G”. Os resultados, exibidos na tabelas 22, 23, 25 e 26 levam às mesmas conclusões obtidas com o experimento usando óleo do tipo “D”. Acrescento, contudo, que estes tipos de óleo tendem a exibir uma perda maior, devido à suas elevadas condutividades. Tal perda maior pode ser observada pelos valores elevados das partes imaginárias de κ_z nas tabelas 22 e 23.

Também foram realizados experimentos em baixas frequências. Em teoria, frequências baixas reduzem a atenuação porque geralmente implicam em menos perdas por efeito Joule.

A frequência escolhida foi de 1 kHz. Nesta frequência, a malha utilizada

Tabela 25: Erro de cada um dos métodos para obtenção dos modos apresentados na tabela 22.

Modo	m	semi-analítico	proposto	mdf
1	1	1.750×10^{-4}	1.649×10^{-4}	3.450×10^{-4}
2	2	1.420×10^{-4}	1.393×10^{-4}	2.263×10^{-4}
3	3	1.407×10^{-4}	1.285×10^{-4}	2.754×10^{-4}
4	0	1.705×10^{-4}	1.282×10^{-4}	3.724×10^{-4}
5	1	2.751×10^{-4}	1.980×10^{-4}	3.908×10^{-4}
6	2	2.605×10^{-4}	2.233×10^{-4}	4.958×10^{-4}
7	3	2.368×10^{-4}	2.275×10^{-4}	6.661×10^{-4}
8	0	2.897×10^{-4}	2.640×10^{-4}	4.974×10^{-4}
9	0	4.172×10^{-4}	3.200×10^{-4}	7.750×10^{-4}

Tabela 26: Erro de cada um dos métodos para obtenção dos modos apresentados na tabela 23.

Modo	m	semi-analítico	proposto	mdf
1	1	1.655×10^{-4}	1.601×10^{-4}	4.000×10^{-4}
2	2	1.674×10^{-4}	1.441×10^{-4}	2.010×10^{-4}
3	3	1.667×10^{-4}	1.270×10^{-4}	2.937×10^{-4}
4	0	1.311×10^{-4}	1.345×10^{-4}	4.782×10^{-4}
5	1	2.473×10^{-4}	1.824×10^{-4}	3.657×10^{-4}
6	2	2.877×10^{-4}	2.900×10^{-4}	4.712×10^{-4}
7	3	2.578×10^{-4}	2.333×10^{-4}	5.920×10^{-4}
8	0	3.093×10^{-4}	2.992×10^{-4}	5.737×10^{-4}
9	0	4.227×10^{-4}	3.811×10^{-4}	5.006×10^{-4}

pele Método das Diferenças Finitas foi truncada para $\rho > 600.000\text{m}$, o que por si só requer uma quantidade muito elevada de pontos e eleva o tempo de execução do Método das Diferenças Finitas consideravelmente. Os outros métodos, inclusive o proposto neste trabalho, sofrem menos com o aumento do comprimento de onda, uma vez que não necessitam de truncamento do domínio.

As tabelas 27 a 30 mostram os primeiros modos para os óleos tipos “D” e “F”. Observa-se que o Método das Diferenças Finitas continua sendo o pior método em termos de acurácia, com erros ainda mais distantes daqueles visualizados no experimento a 0,8 GHz. Além disso, neste caso, o método proposto baseado em PG é claramente superior em termos de acurácia, pois apresenta os menores erros quando comparado aos outros dois métodos. Em termos eletromagnéticos, nota-se que a parte imaginária reduziu bastante de valor devido à menor condutividade dos materiais em baixas frequências.

Do ponto de vista de atenuação, no óleo do tipo “D” (o menos atenuado), a atenuação por metro é dada por $10 \log(e^{-1.902e^{-4}}) = 8.26 \times 10^{-4}$ dB/m. Isso mostra que a redução da frequência de propagação de fato traz vantagens, pois

Tabela 27: Valor de κ_z para os primeiros modos com $\kappa_z \in [10^{-7} - 150i; 100 - 10^{-7}i]$, frequência 1 kHz, óleo tipo “D”, por ordem crescente de atenuação.

	m	semi-analítico (10^{-4}m^{-1})	proposto (10^{-4}m^{-1})	mdf (10^{-4}m^{-1})
1	1	$0.1971 - 1.902i$	$0.1983 - 1.988i$	$0.2562 - 2.007i$
2	2	$0.2231 - 2.491i$	$0.2198 - 2.470i$	$0.3071 - 2.191i$
3	3	$0.2593 - 2.999i$	$0.2619 - 3.005i$	$0.5944 - 6.351i$
4	0	$0.3206 - 3.090i$	$0.3181 - 3.112i$	$0.7059 - 16.001i$
5	1	$0.4428 - 4.222i$	$0.4466 - 4.302i$	$0.6602 - 29.624i$

Tabela 28: Valor de κ_z para os primeiros modos com $\kappa_z \in [10^{-7} - 150i; 100 - 10^{-7}i]$, frequência 1 kHz, óleo tipo “F”, por ordem crescente de atenuação.

	m	semi-analítico (10^{-4}m^{-1})	proposto (10^{-4}m^{-1})	mdf (10^{-4}m^{-1})
1	1	$0.1704 - 6.2330i$	$0.1721 - 6.2218i$	$0.1637 - 7.7742i$
2	2	$0.2010 - 6.7720i$	$0.2025 - 6.7101i$	$0.1977 - 6.8092i$
3	3	$0.2479 - 7.7122i$	$0.2440 - 7.6301i$	$0.2119 - 8.6481i$
4	0	$0.3224 - 11.1077i$	$0.3166 - 11.1239i$	$0.2954 - 15.6888i$
5	1	$0.4513 - 17.815i$	$0.4551 - 17.677i$	$0.5800 - 22.444i$

Tabela 29: Erro de cada um dos métodos para obtenção dos modos apresentados na tabela 27.

Modo	m	semi-analítico	proposto	mdf
1	1	1.331×10^{-3}	1.010×10^{-3}	41.220×10^{-3}
2	2	6.752×10^{-3}	1.997×10^{-3}	37.988×10^{-3}
3	3	7.401×10^{-3}	2.675×10^{-3}	50.783×10^{-3}
4	0	5.572×10^{-3}	3.001×10^{-3}	92.453×10^{-3}
5	1	8.633×10^{-3}	2.081×10^{-3}	110.200×10^{-3}

Tabela 30: Erro de cada um dos métodos para obtenção dos modos apresentados na tabela 28.

Modo	m	semi-analítico	proposto	mdf
1	1	5.984×10^{-3}	2.944×10^{-3}	12.547×10^{-3}
2	2	4.100×10^{-3}	4.218×10^{-3}	11.677×10^{-3}
3	3	3.178×10^{-3}	3.004×10^{-3}	9.001×10^{-3}
4	0	9.753×10^{-3}	8.186×10^{-3}	19.887×10^{-3}
5	1	11.9935×10^{-3}	9.022×10^{-3}	13.943×10^{-3}

implica em redução considerável da atenuação por metro.

Portanto, resultados desta seção mostram que o modelo proposto é capaz de encontrar os modos de propagação com acurácia semelhante ou superior ao modelo semi-analítico e ao modelo baseado em Método das Diferenças Finitas. No entanto, é preciso notar que *o método semi-analítico é mais difícil de codificar*, especialmente devido à presença de integrais envolvendo funções de Bessel no plano complexo que têm erros de truncamento consideráveis para valores grandes do argumento. Mesmo aplicando a modificação 5-46, a implementação numérica ainda é complexa, pois é comum que a integral de linha proveniente do princípio do argumento não convirja, impossibilitando calcular o número de raízes num retângulo. Neste caso, o método subdivide o retângulo em outros 4 e tenta novamente para cada sub-retângulo. Ocasionalmente, o método de busca de raízes precisa subdividir recursivamente muitas vezes para conseguir calcular a integral, o que gera uma árvore de busca muito grande e torna o processo muito lento. Adicionalmente, o modelo proposto também apresenta superioridade de acurácia quando comparado ao Método das Diferenças Finitas. Isso se deve ao fato deste necessitar truncar o domínio para ρ suficientemente grande.

Tais problemas *não ocorrem no modelo proposto baseado em programação genética*, pois este é robusto a indivíduos difíceis de avaliar (eles passam a ter avaliação ruim e saem da população). Outra vantagem do modelo proposto neste trabalho é o ganho de desempenho devido à soluções parametrizadas. Esse aspecto será analisado na próxima seção.

5.5.2 Experimento paramétrico

Nesta seção, serão apresentados os resultados do experimento paramétrico, no qual a técnica proposta resolve o problema de autovalores 5-34 sujeito a variação de dois parâmetros: raio do duto de produção e percentual de incrustação. Em particular, o raio do duto pôde variar na faixa [5; 10] cm e o percentual de incrustação pôde variar na faixa [10%; 25%]. Este é o primeiro experimento nesta tese que apresenta uma solução paramétrica com dois parâmetros.

A título de comparação, o método semi-analítico e o Método das Diferenças Finitas foram empregados para Q^2 combinações de valores dos parâmetros, onde Q é o número de valores adotados para cada parâmetro. Note que os referidos métodos não são paramétricos, ou seja, para avaliar Q^2 combinações, é necessário executar cada um dos métodos Q^2 vezes. O gráfico 13 mostra o tempo de execução acumulado de cada um dos métodos em função de Q . Os experimentos foram executados na frequência 0,8 GHz.

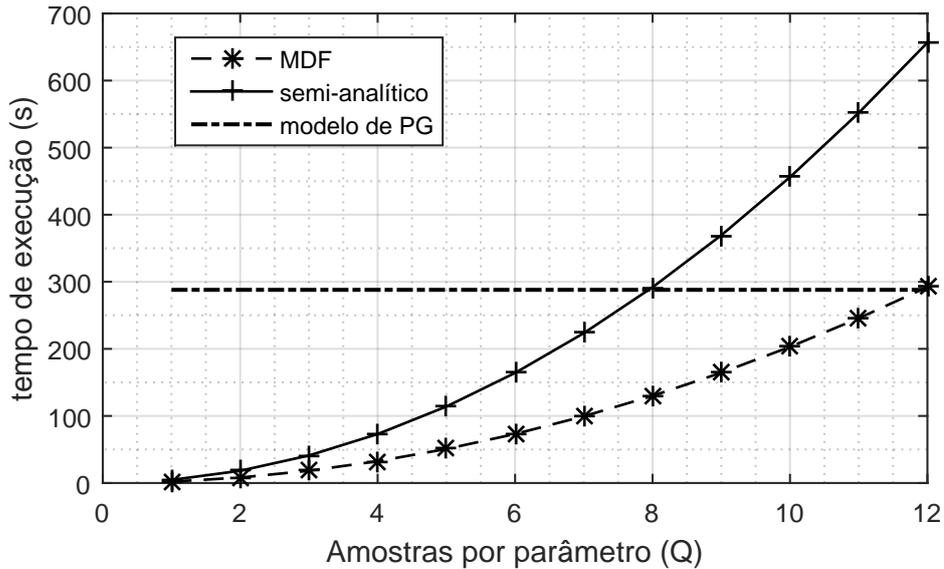


Figura 13: Comparação do tempo de 1 execução do método proposto baseado em PG com Q^2 execuções dos métodos clássicos para determinar o primeiro modo de propagação ($m = 1$) do poço de petróleo na frequência $f = 0,8$ GHz e óleo do tipo “D”.

O gráfico 13 revela que 1 execução do modelo proposto leva próximo de 300s. Este tempo também é o tempo de determinação do primeiro modo, uma vez que, dada a autofunção parametrizada, é possível calcular instantaneamente o valor numérico de κ_z para qualquer combinação de raio do duto de produção e percentual de incrustação na faixa [5; 10] cm para o raio e [10%; 25%] para a incrustação. Adicionalmente, o gráfico também exhibe os tempos acumulados de Q^2 execuções do Método das Diferenças Finitas e do método semi-analítico. Fica evidente que, para $Q = 8$, o tempo de 1 execução do modelo proposto passa a ser menor que o tempo de execução do método semi-analítico, logo, a partir de 64 cenários, é mais eficiente em termos de tempo de execução utilizar o modelo analítico proposto nesta tese. Além disso, um raciocínio análogo pode ser realizado para comparar o modelo proposto e o Método das Diferenças Finitas: a partir de 144 cenários, é mais eficiente em termos de tempo de execução utilizar o modelo analítico proposto nesta tese. Resta saber se o ganho em tempo de execução da solução paramétrica não implica em prejuízo na acurácia.

Para avaliar e comparar a acurácia da técnica proposta com aquela das técnicas clássicas, foi calculado o erro $\epsilon = \frac{\|A_0\mathbf{H} + \kappa_z A_1\mathbf{H} + \kappa_z^2 A_2\mathbf{H}\|}{\|\mathbf{H}\|}$ de 144 execuções de cada técnica clássica utilizando todas as combinações de 12 valores de incrustação e 12 valores de raio do duto de produção. Então, esses erros foram comparados com os erros obtidos a partir da solução paramétrica

Tabela 31: Erros médio, máximo, mínimo e desvio padrão do erro de cada um dos métodos para obtenção do primeiro modo – problema paramétrico para frequência de 0,8 GHz e óleo do tipo “D”.

	analítico baseado em PG	semi-analítico	mdf
Média	$2,213 \times 10^{-4}$	$1,774 \times 10^{-4}$	$3,677 \times 10^{-4}$
Máximo	$23,396 \times 10^{-4}$	$7,600 \times 10^{-4}$	$14,324 \times 10^{-4}$
Mínimo	$0,676 \times 10^{-4}$	$0,0923 \times 10^{-4}$	$1,337 \times 10^{-4}$
Desvio	$0,771 \times 10^{-4}$	$1,101 \times 10^{-4}$	$1,802 \times 10^{-4}$

Tabela 32: Erros médio, máximo, mínimo e desvio padrão do erro de cada um dos métodos para obtenção do primeiro modo – problema paramétrico para frequência de 1 kHz e óleo do tipo “D”.

	analítico baseado em PG	semi-analítico	mdf
Média	$7,701 \times 10^{-3}$	$8,392 \times 10^{-3}$	$13,661 \times 10^{-3}$
Máximo	$31,005 \times 10^{-3}$	$13,934 \times 10^{-3}$	$25,712 \times 10^{-3}$
Mínimo	$1,799 \times 10^{-3}$	$0,902 \times 10^{-3}$	$2,486 \times 10^{-3}$
Desvio	$3,020 \times 10^{-3}$	$3,261 \times 10^{-3}$	$5,772 \times 10^{-3}$

para os mesmos 144 valores dos parâmetros. A tabela tabela 31 resume os resultados.

Os resultados apresentados na tabela 31 mostram que a solução analítica paramétrica entregue pelo modelo proposto apresenta erro médio menor do que aquele apresentado pelo Método das Diferenças Finitas. Além disso, o erro mínimo também é menor, o que mostra que a solução paramétrica apresenta acurácia melhor do que o Método das Diferenças Finitas. Como a única vantagem do Método das Diferenças Finitas é o tempo de execução, conclui-se que, se o número de cenários for grande, é mais vantajoso utilizar o modelo proposto do que o Método das Diferenças Finitas, pois o tempo de execução acumulado tende a ser menor (figura 13) com uma acurácia em média melhor.

Ao comparar com a técnica semi-analítica, a acurácia da solução analítica paramétrica é em média pior, mas ainda assim o erro médio é próximo ao erro médio obtido pelo modelo semi-analítico. Por essa razão, conclui-se que a técnica proposta foi capaz de gerar uma solução paramétrica com boa acurácia e, por isso, é mais adequada para analisar muitos cenários dos parâmetros, pois mais uma vez o tempo de execução acumulado tende a ser menor (ver figura 13).

Os mesmos experimentos foram realizados na frequência de 1 kHz. A figura 14 compara os tempos de execução enquanto que a tabela 32 mostra os erros obtidos.

As conclusões são semelhantes ao experimento na frequência de 0,8 GHz. Note, contudo, que todos os três métodos tiveram um aumento considerável no

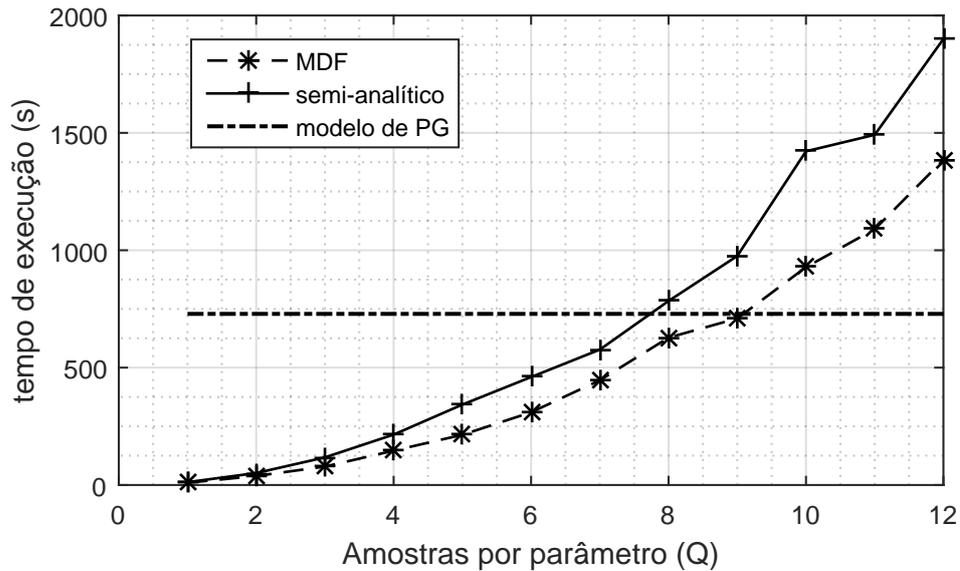


Figura 14: Comparação do tempo de 1 execução do método proposto baseado em PG com Q^2 execuções dos métodos clássicos para determinar o primeiro modo de propagação ($m = 1$) do poço de petróleo na frequência $f = 1$ kHz e óleo do tipo “D”.

tempo de execução e uma redução na acurácia. Isso ocorre devido ao aumento significativo no comprimento de onda, o que cria a necessidade de avaliar o domínio em mais pontos. No caso do Método das Diferenças Finitas, deve-se truncar o domínio para um valor consideravelmente maior de ρ , o que gera uma discretização com muito mais pontos. No caso dos métodos analíticos, o comprimento de onda maior gera funções com decaimento mais lento (lembrese: as funções u e v são nulas no infinito), o que dificulta a convergência de integrais numéricas porque passam a necessitar avaliar a função em mais pontos.

Por fim, a equação 5-47 mostra a melhor expressão encontrada pelo modelo proposto para as funções u e v , com $m = 1$, em função do raio do duto de produção, representado pela variável a , e parametrizado pelo percentual de

Tabela 33: Valores numéricos da parte real dos coeficientes da equação 5-47. Coeficientes menores do que 10^{-7} foram considerados nulos para simplificar a exibição na tabela. O experimento usou óleo tipo “D” e frequência de 1 kHz.

j	A_j	B_j	C_j	D_j	E_j	F_j
1	1,212	$2,948 \times 10^{-6}$	0	$5,054 \times 10^{-3}$	$3,017 \times 10^{-5}$	0
2	2,945	8,331	0	$2,310 \times 10^{-3}$	$4,706 \times 10^{-3}$	0
3	6,881	7,001	0	$3,556 \times 10^{-3}$	$8,196 \times 10^{-3}$	0
4	2,945	1,431	0	$8,081 \times 10^{-3}$	$0,918 \times 10^{-3}$	0
5	0,2201	4,700	0	$2,333 \times 10^{-3}$	$1,399 \times 10^{-3}$	0
6	0,0167	6,405	0	$1,914 \times 10^{-3}$	$0,508 \times 10^{-3}$	0
7	0	0	$1,811 \times 10^5$	0	0	$2,618 \times 10^3$

Tabela 34: Expressão analítica dos coeficientes κ_ρ^j da equação 5-47. O experimento usou óleo tipo “D” e frequência de 1 kHz.

j	κ_ρ^j
1	$0,97\kappa_1^2 - (0.196 - 1.98i) - (0.220 - 2.02i)(1, 1 - p)a - (0.269 - 2.40i)a^2$
2	$1,00\kappa_2^2 - (0.198 - 2.00i) - (0.221 - 2.01i)(1, 1 - p)a - (0.271 - 2.41i)a^2$
3	$1,24\kappa_3^2 - (0.201 - 1.98i) - (0.216 - 2.25i)(1, 1 - p)a - (0.267 - 2.39i)a^2$
4	$1,00\kappa_4^2 - (0.189 - 1.96i) - (0.219 - 2.02i)(1, 1 - p)a - (0.271 - 2.40i)a^2$
5	$0,96\kappa_5^2 - (0.198 - 1.98i) - (0.220 - 2.02i)(1, 0 - p)a - (0.270 - 2.40i)a^2$
6	$1,07\kappa_6^2 - (0.207 - 1.99i) - (0.222 - 2.02i)(1, 1 - p)a - (0.270 - 2.39i)a^2$
7	$0,99\kappa_7^2 - (0.200 - 1.98i) - (0.221 - 2.02i)(1, 1 - p)a - (0.272 - 2.40i)a^2$

incrustação, representado por p . A solução foi obtida em 1 kHz.

$$\begin{aligned}
 \hat{u}_j(\rho) &= A_j J_1(\kappa_\rho^j \rho) \\
 &+ B_j Y_1(\kappa_\rho^j \rho) \\
 &+ C_j H_1^{(2)}(\kappa_\rho^j \rho) && \rho_{j-1} \leq \rho < \rho_j \\
 \hat{v}_j(\rho) &= D_j J_1(\kappa_\rho^j \rho) \\
 &+ E_j Y_1(\kappa_\rho^j \rho) \\
 &+ F_j H_1^{(2)}(\kappa_\rho^j \rho) && \rho_{j-1} \leq \rho < \rho_j \\
 \kappa_\rho^j &= \text{função de } a \text{ e } p \\
 j &= 1, 2, \dots, 7.
 \end{aligned} \tag{5-47}$$

A tabela 33 exhibe os valores numéricos dos coeficientes para cada meio de propagação. Além disso, a tabela 34 exhibe as expressões de κ_ρ^j para cada meio de propagação. Note que há uma semelhança grande entre as expressões, uma vez que $\kappa_\rho^2 = \kappa^2 - \kappa_z^2$ e κ_z é igual, por hipótese, em todos os meios.

Como conclusão da seção, considero que o modelo proposto é capaz de gerar funções parametrizadas pelo raio do duto de produção e pelo percentual de incrustação com boa acurácia para valores do parâmetro dentro de uma

faixa pré-especificada. Em particular, neste experimento, o raio foi limitado ao intervalo [5; 10] cm e a incrustação foi limitada ao intervalo [10%; 25%] do valor do raio. Além disso, acrescento que a função paramétrica gerada pelo modelo proposto foi capaz de prever o primeiro autopar com acurácia semelhante àquela obtida pelos métodos clássicos. Contudo, é preciso enfatizar que os métodos clássicos necessitam ser executados para cada combinação de valor do raio do duto e do percentual de incrustação, enquanto que o modelo proposto não precisa, uma vez que a solução paramétrica pode ser reutilizada. Sendo assim, para uma quantidade grande de cenários dos parâmetros, é mais eficiente em termos de tempo de execução acumulado utilizar a técnica proposta, como mostram as figuras 13 e 14.

5.6

Conclusões do Capítulo

Neste capítulo, foi analisado o problema de propagação de ondas eletromagnéticas em poços de petróleo. Foram adotadas duas óticas: propagação em altas frequências (em particular 0,8 GHz) e em baixas frequências (1 kHz). Em altas frequências, o sinal fica confinado no duto de produção, mas há muitas perdas por efeito Joule. Por outro lado, em baixas frequências, não há perdas significativas por efeito Joule, mas o sinal vaza para fora do poço. Portanto, não é óbvio qual faixa de frequências geraria menos perdas.

O problema de propagação de ondas no interior do duto de produção foi formulado como um problema de determinação de modos em um guia de ondas radialmente fragmentado, contendo os vários materiais tipicamente encontrados em um poço. Este problema, por sua vez, foi escrito como um problema de autovalores quadrático, com operadores não-normais. A técnica proposta foi posta a encontrar os autovalores com menor parte imaginária (mais próxima de 0) e os resultados foram comparados com o método semi-analítico baseado no princípio do argumento e com o Método das Diferenças Finitas, por ser uma técnica puramente numérica.

Em termos de acurácia, a técnica proposta apresentou semelhança com a técnica semi-analítica, sendo melhor em baixa frequência. Além disso, a técnica proposta apresentou acurácia superior ao Método das Diferenças Finitas em todos os experimentos, sobretudo em baixas frequências. Isso ocorre porque a malha de pontos utilizada pelo Método das Diferenças Finitas e pelo Método dos Elementos Finitos não é precisa devido ao truncamento do domínio infinito.

Em termos de tempo de execução, a técnica proposta aparenta ser mais lenta do que as técnicas utilizadas como comparação, devido ao processamento simbólico e à presença de uma evolução por programação genética. Contudo,

essa aparente desvantagem torna-se uma vantagem quando deseja-se analisar vários cenários do mesmo problema. Neste caso, a solução paramétrica, mesmo levando mais tempo para ser obtida, pode ser reutilizada indefinidamente para diversos cenários de valores dos parâmetros, enquanto as técnicas clássicas precisam ser re-executadas para cada valor particular dos parâmetros. Essa característica foi explicitada através dos experimentos paramétricos utilizando o raio do duto de produção e o percentual de incrustação como parâmetros. Os experimentos realizados mostraram que o tempo de execução acumulado de 64 execuções do método semi-analítico e 144 execuções do Método das Diferenças Finitas é maior do que o tempo de determinação da solução paramétrica, tanto em alta quanto em baixa frequências. Isso mostra que, para um número grande de cenários, é mais vantajoso em termos de tempo de execução utilizar a técnica proposta.

Ao comparar a acurácia da solução analítica paramétrica com aquela dos métodos clássicos, foi observado que a solução paramétrica apresenta acurácia semelhante ao método semi-analítico e superior ao Método das Diferenças Finitas, sobretudo em baixa frequência, onde este não desempenha bem. Portanto, conclui-se que o método proposto é o mais indicado para problemas de autovalores paramétricos nos quais deseja-se obter a solução para um número grande de valores dos parâmetros.

Do ponto de vista eletromagnético, os experimentos mostraram que o duto de produção tem alta perda por efeito Joule, sobretudo em alta frequência. Em baixa frequência, a perda por efeito Joule é menor, mas, em compensação, há muito vazamento de sinal pelas paredes do poço. No entanto, mesmo com o vazamento de sinal, as perdas são menores em baixa frequência, sendo esta, portanto, mais adequada para montar um enlace.

Com relação aos parâmetros, conclui-se que a incrustação, por apresentar baixa condutividade elétrica, não prejudica a propagação. Portanto, quanto maior o percentual de incrustação, menos perdas. Contudo, note que a incrustação é um processo químico fora do controle do engenheiro. O raio do duto de produção, por outro lado, tem influência direta na atenuação dos modos: quanto maior, menos atenuação.

No próximo capítulo, serão apresentadas as conclusões finais deste trabalho.

6

Conclusões e Trabalhos Futuros

Este capítulo encerra o texto desta tese respondendo à pergunta: quais são as conclusões e os próximos passos desta pesquisa? Para isso, será feito um breve resumo do exposto e, na sequência, apresento as conclusões finais.

Este texto apresentou um novo modelo de solução de problemas de autovalores de operadores diferenciais lineares. Este modelo é assistido por programação genética e trabalha com as expressões analíticas das autofunções, reduzindo ao máximo a quantidade de computações numéricas.

O capítulo 2 mostrou a formulação matemática do problema de autovalores de operadores diferenciais lineares. Além disso, o capítulo fez um resumo de técnicas numéricas geralmente empregadas na literatura para resolver tais problemas, dentre elas o Método dos Elementos Finitos e o Método das Diferenças Finitas. Na sequência, o capítulo apresentou uma revisão literária de trabalhos semelhantes ao presente trabalho, utilizando Programação Genética. Concluiu-se que não há trabalhos na literatura empregando Programação Genética para resolver o problema de autovalores de operadores diferenciais. Por isso, a proposta deste trabalho é uma contribuição original.

O capítulo 3 apresentou o modelo de solução de problemas de autovalores de operadores diferenciais assistido por programação genética. Neste capítulo, foi feita uma descrição da técnica de programação genética, com cromossomos baseados em árvore. Na sequência, o capítulo propôs uma forma de representar funções vetoriais em termos de combinações lineares de árvores, possivelmente com superposição de árvores. Depois, o capítulo mostrou como formular um problema de otimização de um funcional de modo que a solução seja a autofunção desejada. Posteriormente, o capítulo descreveu como “remover” as autofunções já encontradas para que uma nova otimização possa buscar por outras autofunções. A técnica foi desenvolvida de modo que as autofunções possam ser parametrizadas por constantes do problema, como diâmetros, comprimentos e constantes físicas. No final do capítulo, foi apresentada uma adaptação ao modelo para que este possa resolver problemas de autovalores de operadores não-normais.

Os capítulos 4 e 5 mostram estudos de caso de aplicação do modelo proposto. No capítulo 4, são apresentados os casos de determinação de auto-

funções do oscilador harmônico quântico e do guia de ondas retangular. Estes exemplos foram escolhidos por apresentarem solução analítica simples e, com isso, servem de prova de conceito. Os resultados mostraram que o modelo proposto conseguiu encontrar as autofunções conforme esperado na maioria das rodadas. Em particular, o modelo paramétrico apresentou resultado preliminar positivo para resolver o problema do guia de ondas retangular, onde o parâmetro era uma constante que representava o tamanho da seção reta do guia. Neste capítulo, contudo, os problemas estudados envolviam apenas operadores normais.

Ainda no capítulo 4, foi apresentado um estudo de caso cuja função era estressar ao máximo as técnicas. Para isso, utilizou-se como base o problema de autopares do operador de Laplace, mas com domínio pseudo-retangular, onde uma das arestas fora substituída por uma senóide. A frequência desta senóide foi controlada de modo a analisar a acurácia de cada técnica. Os resultados mostraram que a técnica proposta é pouco influenciada pela presença da senóide no domínio, em oposição ao Método das Diferenças Finitas e ao Método dos Elementos Finitos, pois estes apresentaram perda de acurácia. Esta se deve à dificuldade de modelar discretamente um domínio com variações geométricas de alta frequência. Como a técnica proposta é analítica, ela não sofre com variações bruscas na geometria do domínio, o que a confere melhor performance neste experimento.

O capítulo 5 apresentou um estudo de caso do problema de propagação de ondas eletromagnéticas em poços de petróleo. O capítulo formulou um problema de autovalores quadrático e mostrou que este não se trata de um problema envolvendo operadores auto-adjuntos ou normais. Portanto, a solução demandou a versão do modelo adequada à operadores que não são normais ou auto-adjuntos. O capítulo apresentou um estudo de caso para um único poço de geometria e materiais típicos, e confrontou o resultado do modelo proposto com aqueles obtidos por um segundo modelo, semi-analítico. Com isso, este trabalho mostrou que o modelo proposto no capítulo 3 é capaz de encontrar soluções próximas àsquelas obtidas pelo modelo semi-analítico tradicional, com a vantagem de apresentar menos dificuldades numéricas.

Em termos de acurácia, o modelo proposto apresentou acurácia semelhante ou melhor do que os métodos usados como comparação para determinação dos primeiros modos de propagação para três tipos diferentes de óleo. Mais especificamente, o método proposto teve acurácia melhor do que o Método das Diferenças Finitas, sobretudo em baixas frequências. Além disso, a acurácia do método proposto foi semelhante àquela do método semi-analítico, sendo superior em alguns casos, principalmente em baixas frequências.

Adicionalmente, ainda no capítulo 5, os experimentos paramétricos mostraram que o tempo de execução da técnica proposta para obtenção da solução paramétrica é menor do que 64 execuções do modelo semi-analítico e 144 execuções do Método das Diferenças Finitas, ou seja, para avaliação de muitos cenários, é mais interessante utilizar a técnica proposta para obter uma solução paramétrica. Adicionalmente, o capítulo 5 mostrou que a solução paramétrica apresentou acurácia semelhante à solução obtida pelos métodos comparativos para valores particulares dos parâmetros, o que reforça o argumento de utilizar o método proposto em problema nos quais deseja-se avaliar múltiplos cenários.

Como conclusões finais, destaco a capacidade do modelo proposto de encontrar soluções comparáveis ou melhores do que aquelas obtidas pelos modelos numéricos e semi-analíticos clássicos. Para um número de gerações suficientemente grande, o modelo proposto sempre converge para uma solução com erro arbitrariamente pequeno, o que é uma característica desejável de qualquer método computacional.

Outra conclusão é que o modelo proposto é capaz de gerar soluções analíticas parametrizadas. Estas soluções, conhecidas como *proxies*, aproximam a autofunção e os autovalores para diferentes valores dos parâmetros. Neste caso, há um ganho de desempenho em relação às técnicas tradicionais, que demandam ser executadas novamente para cada combinação de parâmetros que se deseja avaliar. Sendo assim, o modelo proposto é adequado para problemas que envolvem avaliação de muitos cenários, pois a tendência é que haja redução no tempo total de execução, quando comparado à outras técnicas.

Como movimentos futuros dessa pesquisa, destaco a necessidade de implementar o modelo proposto utilizando uma linguagem de programação de alta performance, como C++, de modo a reduzir os tempos de execução. Outra opção é utilizar evolução em código de máquina, que é uma vertente da Programação Genética que trabalha com cromossomos lineares. Como as funções são geradas em código de máquina, suas avaliações são muito rápidas, o que dispensaria simplificação simbólica e, em teoria, poderia aumentar muito o desempenho do método proposto.

Além disso, do ponto de vista científico, é interessante aplicar a técnica para outras áreas, além da Física, que utilizem equações diferenciais parciais e problemas de autovalores. Neste caso, destaco as áreas relativas à Econofísica, que trabalham com modelos de Física aplicados à Macroeconomia [67]. Adicionalmente, acrescento a área de propagação de fluidos em meios porosos, como campos de petróleo [68]. Estes costumam apresentar muita incerteza e, assim, demandam centenas de simulações para estimar o valor presente líquido esperado.

Do ponto de vista da técnica proposta, é interessante generaliza-la para operadores integro-diferenciais, que são operadores definidos em termos de derivadas e integrais. Sendo assim, alguns problemas importantes da Física passam a ser abordados pelo modelo proposto, como o problema de espalhamento e cálculo de seção de radar em Eletromagnetismo [59, 69].

Referências bibliográficas

- 1 HUTTON, D. *Fundamentals of finite element analysis*. [S.l.]: McGraw-Hill,, 2004.
- 2 MOAVENI, S. *Finite Element Analysis Theory and Application with ANSYS*, 3/e. [S.l.]: Pearson Education India, 2008.
- 3 JIN, J.-M. *The finite element method in electromagnetics*. [S.l.]: John Wiley & Sons, 2015.
- 4 HANSON, A. B. Y. G. W. *Operator theory for electromagnetics: an introduction*. [S.l.]: Springer Science+Business Media New York, 2002.
- 5 NAYLOR, G. R. S. A. W. *Linear Operator Theory in Engineering and Science*. [S.l.]: Springer Science & Business Media, 2000.
- 6 HSING, R. E. T. *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. [S.l.]: Wiley, 2015. (Wiley Series in Probability and Statistics).
- 7 GROSSMANN CHRISTIAN, R. H.-G. S. M. *Numerical Treatment of Partial Differential Equations*. [S.l.]: Springer Science & Business Media, 2007.
- 8 ISERLES, A. *A First Course in the Numerical Analysis of Differential Equations*. Second. Cambridge University Press, 2008. Cambridge Books Online. ISBN 9780511995569. Disponível em: <<http://dx.doi.org/10.1017/CBO9780511995569>>.
- 9 LEVEQUE, R. J. *Finite Volume Methods for Hyperbolic Problems*. [S.l.]: Cambridge University Press, 2002.
- 10 HARRINGTON, R. F. *Field Computation by Moment Methods*. [S.l.]: IEEE Press, 1968.
- 11 BANERJEE, P. K. *Boundary Element Methods in Engineering and Sciences*. [S.l.]: McGraw-Hill, 1994.
- 12 GOTTLIEB, D.; ORZAG, S. *Numerical Analysis of Spectral Methods: Theory and Applications*. [S.l.]: Society for Industrial and Applied Mathematics, 1977.
- 13 HOFFMAN, J. D. *Numerical Methods for Engineers and Scientists*. [S.l.]: McGraw-Hill, Inc., 1992.
- 14 SMITH, G. D. *Numerical Solution of Partial Differential Equations: Finite Difference Methods (Oxford Applied Mathematics and Computing Science Series)*. [S.l.]: Clarendon Press, 1986.

- 15 JOHNSON, S. G. *Notes on Perfectly Matched Layers (PMLs)*. [S.l.]: online MIT course notes, 2010.
- 16 BERENGER, J.-P. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, Elsevier, v. 114, n. 2, p. 185–200, 1994.
- 17 SAAD, Y. *Iterative methods for sparse linear systems*. [S.l.]: Society for Industrial and Applied Mathematics, 2003.
- 18 DAVIES, A. J. *The finite element method: a first approach*. [S.l.]: Oxford University Press, 1980.
- 19 SCHINDLER, M.; CHEN, E. *Barycentric Coordinates in Olympiad Geometry*. 2012.
- 20 WEISSTEIN, E. W. Barycentric coordinates. Wolfram Research, Inc., 2003.
- 21 COMSOL. *COMSOL Multiphysics®*. 2016. <<https://br.comsol.com/>>. Acessado: 26-07-2016.
- 22 TECHNOLOGIES, C. S. *CST® Microwave Studio®*. 2016. <<https://www.cst.com/products/cstmws>>. Acessado: 26-07-2016.
- 23 SYSTEMES, B. D. *Solidworks Simulation Premium*. 2016. <<http://www.solidworks.com/sw/products/simulation/simulation-premium.htm>>. Acessado: 26-07-2016.
- 24 KOZA, J. R. *Genetic programming: on the programming of computers by means of natural selection*. Cambridge, Mass. MIT Press, 1992. (Complex adaptive systems). A Bradford book. ISBN 0-262-11170-5. Disponível em: <<http://opac.inria.fr/record=b1082356>>.
- 25 LOBÃO, W. J. de A. *Solução de Equações Diferenciais Ordinárias, Parciais e Estocásticas por Programação Genética e Diferenciação Automática*. Tese (Doutorado) — PUC-Rio, 2015.
- 26 IMAE, J. et al. Design of nonlinear control systems by means of differential genetic programming. In: IEEE. *Decision and Control, 2004. CDC. 43rd IEEE Conference on*. [S.l.], 2004. v. 3, p. 2734–2739.
- 27 TSOULOS, I. G.; LAGARIS, I. E. Solving differential equations with genetic programming. *Genetic Programming and Evolvable Machines*, Springer, v. 7, n. 1, p. 33–54, 2006.
- 28 DARWIN, C. *A origem das Espécies e a Seleção Natural*. [S.l.]: Ed. Hemus, 2000.
- 29 MCCARTHY, J. Recursive functions of symbolic expressions and their computation by machine, part i. *Communications of the ACM*, ACM, v. 3, n. 4, p. 184–195, 1960.
- 30 BÖHM, W.; GEYER-SCHULZ, A. Exact uniform initialization for genetic programming. In: *FOGA*. [S.l.: s.n.], 1996. p. 379–407.

- 31 CHELLAPILLA, K. Evolutionary programming with tree mutations: Evolving computer programs without crossover. *Genetic Programming*, p. 431–438, 1997.
- 32 HINCHLIFFE, M. et al. Modelling chemical process systems using a multi-gene. *Late Breaking Papers at the Genetic Programming*, p. 56–65, 1996.
- 33 CHERKASSKY, V.; GEHRING, D.; MULIER, F. Comparison of adaptive methods for function estimation from samples. *IEEE Transactions on Neural Networks*, IEEE, v. 7, n. 4, p. 969–984, 1996.
- 34 BROYDEN, C. G. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA Journal of Applied Mathematics*, IMA, v. 6, n. 1, p. 76–90, 1970.
- 35 SHANNO, D. F. Conditioning of quasi-newton methods for function minimization. *Mathematics of computation*, v. 24, n. 111, p. 647–656, 1970.
- 36 COLEMAN, T. F.; LI, Y. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on optimization*, SIAM, v. 6, n. 2, p. 418–445, 1996.
- 37 SHAMPINE, L. F. Vectorized adaptive quadrature in matlab. *Journal of Computational and Applied Mathematics*, Elsevier, v. 211, n. 2, p. 131–140, 2008.
- 38 READ, W. Analytical solutions for a helmholtz equation with dirichlet boundary conditions and arbitrary boundaries. *Mathematical and computer modelling*, Elsevier, v. 24, n. 2, p. 23–34, 1996.
- 39 MCCARTIN, B. J. *Laplacian eigenstructure of the equilateral triangle*. [S.l.]: Hikari, Limited, 2011.
- 40 ABRAMOWITZ, M.; STEGUN, I. A.; ROMER, R. H. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. [S.l.]: AAPT, 1988.
- 41 ANDREWS, G. E.; ASKEY, R.; ROY, R. *Special functions, Encyclopedia of mathematics and its applications, Vol. 71*. [S.l.]: Cambridge University Press, Cambridge, 1999.
- 42 OLVER, F. et al. Nist handbook of mathematical functions. *New York*, Cambridge University Press, 2010.
- 43 VERMEHREN, V.; OLIVEIRA, H. de. Close expressions for meyer wavelet and scale function. *arXiv preprint arXiv:1502.00161*, 2015.
- 44 DAUBECHIES, I. *Ten lectures on wavelets*. [S.l.]: SIAM, 1992.
- 45 BASICS of Wavelets. Autor Desconhecido. <<http://math.bu.edu/people/mkon/Wavelets.pdf>>. Acessado: 08-08-2017.

- 46 LEHOUCQ, R. B.; SORENSEN, D. C. Deflation techniques for an implicitly restarted arnoldi iteration. *SIAM Journal on Matrix Analysis and Applications*, SIAM, v. 17, n. 4, p. 789–821, 1996.
- 47 MATHWORKS. *Subset of eigenvalues and eigenvectors*. <<https://www.mathworks.com/help/matlab/ref/eigs.html>>. Acessado: 08-08-2017.
- 48 MATHWORKS. *Create triangular or tetrahedral mesh*. <<https://www.mathworks.com/help/pde/ug/generatemesh.html>>. Acessado: 08-08-2017.
- 49 QUANTUM Harmonic Oscillator. <<http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/hosc2.html#c1>>. Acessado: 03-08-2016.
- 50 TROFIMENKOFF, F. N. et al. Characterization of em downhole-to-surface communication links. *Geoscience and Remote Sensing, IEEE Transactions on*, IEEE, v. 38, n. 6, p. 2539–2548, 2000.
- 51 WEI, Y. *Propagation of electromagnetic signal along a metal well in an inhomogeneous medium*. Tese (Doutorado) — Norwegian University of Science and Technology, 2013.
- 52 ROSA, G. S.; BERGMANN, J. R.; ZANG, S. R. Electromagnetic analysis of wireless telemetry for oil fields. In: IEEE. *Microwave & Optoelectronics Conference (IMOC), 2013 SBMO/IEEE MTT-S International*. [S.l.], 2013. p. 1–4.
- 53 CUNHA, A.; PACHECO, M. et al. A simulation model for electromagnetic wave propagation within production pipes of oil wells. In: OFFSHORE TECHNOLOGY CONFERENCE. *OTC Brasil*. [S.l.], 2015.
- 54 KAUFMAN, A. A. The electrical field in a borehole with a casing. *Geophysics*, Society of Exploration Geophysicists, v. 55, n. 1, p. 29–38, 1990.
- 55 VIDAL, J. L. A. *Sistema Inteligente De Comunicação e Controle Sem Fio Baseado Em Materiais Com Memória De Forma*. Tese (Doutorado) — COPEE/UFRJ, 2010.
- 56 NASCIMENTO, J. C. S. *Simulador de Escoamento Multifásico em Poços de Petróleo*. Tese (Doutorado) — Universidade Federal do Rio Grande do Norte, 2013.
- 57 MAINIER, F. B.; FREITAS, A. E. R. de; FIGUEIREDO, A. A. M. Acid removal of mineralized incrustation in petroleum production pipe. *International Journal of Research and Reviews in Applied Sciences*, Academic Research Publishing Agency:(arapapress), v. 23, n. 1, p. 65, 2015.
- 58 TAI, C.-T.; IEEE. *Dyadic Green Functions in Electromagnetic Theory (IEEE Press Series on Electromagnetic Waves)*. 2 sub. ed. [S.l.]: Institute of Electrical & Electronics Engineer, 1994. ISBN 9780780304499.
- 59 BALANIS, C. A. *Advanced engineering electromagnetics*. 2. ed. [S.l.]: John Wiley & Sons, Inc., 2012.

- 60 PINCHERLE, L. Electromagnetic waves in metal tubes filled longitudinally with two dielectrics. *Physical Review*, APS, v. 66, n. 5-6, p. 118, 1944.
- 61 WHITES, K. W. *Electromagnetic wave propagation through circular waveguides containing radially inhomogeneous lossy media*. [S.l.], 1989.
- 62 SEMENOV, V. Method for the calculation of all non-multiple zeros of an analytic function. *Computational Methods in Applied Mathematics Comput. Methods Appl. Math.*, v. 11, n. 1, p. 67–74, 2011.
- 63 JOHNSON, T.; TUCKER, W. Enclosing all zeros of an analytic function - a rigorous approach. *Journal of Computational and Applied Mathematics*, Elsevier, v. 228, n. 1, p. 418–423, 2009.
- 64 YOUNG, K.; FREDERIKSE, H. Compilation of the static dielectric constant of inorganic solids. *Journal of Physical and Chemical Reference Data*, AIP Publishing, v. 2, n. 2, p. 313–410, 1973.
- 65 YOUNGLOVE, B.; ELY, J. F. Thermophysical properties of fluids. ii. methane, ethane, propane, isobutane, and normal butane. *Journal of Physical and Chemical Reference Data*, AIP Publishing, v. 16, n. 4, p. 577–798, 1987.
- 66 FOLGERØ, K. Broad-band dielectric spectroscopy of low-permittivity fluids using one measurement cell. *Instrumentation and Measurement, IEEE Transactions on*, IEEE, v. 47, n. 4, p. 881–885, 1998.
- 67 OLKHOV, V. Econophysics macroeconomic model. *arXiv preprint arXiv:1701.06625*, 2017.
- 68 TRANGENSTEIN, J. A.; BELL, J. B. Mathematical structure of the black-oil model for petroleum reservoir simulation. *SIAM Journal on Applied Mathematics*, SIAM, v. 49, n. 3, p. 749–783, 1989.
- 69 NICOLAESCU, L.; OROIAN, T. Radar cross section. In: IEEE. *Telecommunications in Modern Satellite, Cable and Broadcasting Service, 2001. TELSIS 2001. 5th International Conference on*. [S.l.], 2001. v. 1, p. 65–68.

A

Produtos internos

O produto interno entre os vetores a , $a = (a_1, a_2, \dots, a_N)$, e b , $b = (b_1, b_2, \dots, b_N)$, do \mathbb{R}^N é:

$$\langle a; b \rangle = a \cdot b = \sum_{i=1}^N a_i b_i. \quad (\text{A-1})$$

Por outro lado, no caso de a e b serem vetores do \mathbb{C}^N , temos o produto interno:

$$\langle a; b \rangle = a \cdot \bar{b} = \sum_{i=1}^N a_i \bar{b}_i, \quad (\text{A-2})$$

isto é, o elemento à direita precisa ser conjugado.

Para funções, o produto interno é escrito em termos de integrais. Se $f(x)$ e $g(x)$ são funções com domínio em um subconjunto D do \mathbb{R}^N ($D \subseteq \mathbb{R}^N$) e contradomínio no \mathbb{C}^M , então o produto interno é dado por:

$$\langle f; g \rangle_w = \int_D f(x) \cdot \overline{g(x)} w(x) dx, \quad (\text{A-3})$$

onde $w(x)$ é uma função real positiva e o produto escalar é dado pela eq. A-1:

$$f(x) \cdot \overline{g(x)} = \sum_{i=1}^M f_i(x) \overline{g_i(x)}. \quad (\text{A-4})$$

O papel da função $w(x)$ é dar pesos diferentes para diversos valores de $x \in D$.

Todos os produtos internos aqui exemplificados podem ser usados para construir uma definição de norma para o espaço em questão, utilizando

$$\|x\| = \sqrt{\langle x; x \rangle}. \quad (\text{A-5})$$