



Iury Steiner de Oliveira Bezerra

**Investigação de um modelo ab initio para cálculos de
estrutura eletrônica de átomos por Algoritmos Evolutivos,
Wavelets e Polinômios de Laguerre**

Tese de Doutorado

Tese apresentada ao programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da PUC-Rio como parte dos requisitos parciais para obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Prof. Marco Aurélio Cavalcanti Pacheco
Co-orientador: Carlos Kubrusly

Rio de Janeiro
Maio de 2014



Iury Steiner de Oliveira Bezerra

**Investigação de um Modelo ab Initio para Cálculos
de Estrutura Eletrônica de Átomos por Algoritmos
Evolutivos, Wavelets e Polinômios de Laguerre**

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

Prof. Marco Aurélio Cavalcanti Pacheco
Orientador

Departamento de Engenharia Elétrica - PUC-Rio

Prof. Carlos Kubrusly
Co-Orientador

Departamento de Engenharia Elétrica - PUC-Rio

Dr. Douglas Mota Dias

Departamento de Engenharia Elétrica - PUC-Rio

Prof. Alberico Borges Ferreira da Silva
USP

Prof. Mario Norberto Baibich
UFRGS

Prof. André Vargas Abs da Cruz
UEZO

Prof. Renato Portugal
LNCC

Prof. Thiago Messias Cardozo
UFRJ

Prof. José Eugenio Leal
Coordenador Setorial do Centro
Técnico Científico - PUC-Rio

Rio de Janeiro, 13 de maio de 2014

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

Iury Steiner de Oliveira Bezerra

Engenheiro Eletricista pela USU, obteve o título de mestre em Engenharia Elétrica pela Pontifícia Universidade Católica do Rio de Janeiro em 2009, onde também concluiu o Doutorado em Engenharia Elétrica em 2014.

Ficha Catalográfica

Bezerra, Iury Steiner de Oliveira

Investigação de um modelo ab initio para cálculos de estrutura eletrônica de átomos por algoritmos evolutivos, wavelets e polinômios de Laguerre / Iury Steiner de Oliveira Bezerra; orientador: Marco Aurélio Cavalcanti Pacheco; co-orientador: Carlos Kubrusly. – 2014.

147 f.: il. (color.) ; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2014.

Inclui bibliografia

1. Engenharia elétrica – Teses. 2. Nanotecnologia. 3. Química computacional. 4. Química quântica. 5. Mecânica quântica. 6. Wavelets. 7. Algoritmos evolucionários. 8. Algoritmos genéticos. I. Pacheco, Marco Aurélio Cavalcanti. II. Kubrusly, Carlos. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

Agradecimentos

Ao CNPq e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Aos meus orientadores e amigos Marco Aurélio C. Pacheco e Carlos Kubrusly, pelo incondicional apoio durante a realização desse trabalho e todas as lições durante todos esses anos de convivência.

A minha querida professora Marley e a Dilza que também depositaram muita confiança em mim durante a minha passagem no Laboratório de Inteligência Computacional Aplicada.

Ao meu querido professor Geraldo Monteiro Sigaud por todos os ensinamentos passados, sem os quais esses o desenvolvimento desta pesquisa não seria possível.

Aos demais professores pelos ensinamentos.

A minha amiga Dilza, pelo apoio e ensinamentos transferidos durante longos bate-papos.

Aos amigos Eugênio, Alan Bontempo, Alexandre, Daniel, Adriano, Marco Botelho e Harold.

E por fim, agradeço ao meu amigo João Zanni por ter estado disponível e sempre disposto a oferecer sua ajuda.

Aos meus pais meus irmãos e a Mirian, pela educação, atenção e carinho de todas as horas.

Resumo

Bezerra, Iury Steiner de Oliveira; Pacheco, Marco Aurélio Cavalcanti (Orientador); Kubrusly, Carlos (Co-orientador). **Investigação de um modelo ab initio para cálculos de estrutura eletrônica de átomos por Algoritmos Evolutivos, Wavelets e Polinômios de Laguerre**. Rio de Janeiro, 2014. 147p. Tese de Doutorado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

A simulação da estrutura eletrônica de átomos e moléculas, desde do início da década de 90, tem se mostrado uma ferramenta imprescindível para o desenvolvimento de áreas estratégicas, ainda emergentes, mas fundamentais, como por exemplo, a área de Nanotecnologia. No entanto, esse tipo de simulação, ainda hoje é de grande complexidade e exige alto poder computacional. Dessa forma, torna-se fundamental a criação de métodos de simulação mais precisos e computacionalmente menos custosos. Este trabalho utiliza Algoritmos Evolutivos e métodos de Inteligência Computacional junto de algumas ferramentas desenvolvidas e estudadas tradicionalmente pela Matemática Aplicada em cálculos de estrutura eletrônica. Em particular, são construídas novas formas de aproximação de soluções para equação de Schrödinger, que contemplem os requisitos físicos necessários. Essas soluções serão nomeadas de funções de ondas evolucionárias, que neste trabalho serão tratadas como pontos em um espaço de Hilbert formado pelo fecho em L^2 da interseção das funções definidas na semi-reta, contínuas e de derivadas contínuas, simétricas ou antissimétricas com relação à permutação de suas coordenadas. São demonstrados alguns resultados, requisitos para utilização de Algoritmos Evolucionários e Séries de Fourier Generalizadas, baseadas em polinômios de Laguerre modificados e Wavelets. Esta pesquisa é desenvolvida inicialmente para sistemas de dois elétrons, e mais tarde é estendida para sistemas mais complexos, a fim de criar uma abordagem alternativa as tradicionais.

Palavras-chave

Nanotecnologia; Química Computacional; Química Quântica; Mecânica Quântica; Wavelets, Algoritmos Evolucionários; Algoritmos Genéticos.

Abstract

Bezerra, Iury Steiner de Oliveira; Pacheco, Marco Aurélio Cavalcanti (Advisor); Kubrusly, Carlos (Co-advisor). **Investigation of an ab initio model to electronic structure of atoms based on evolutive algoritms, wavelets and Laguerre polynomials**. Rio de Janeiro, 2014. 147p. DSc. Thesis – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

The simulation of the electronic structure of atoms and molecules has been shown to be, from the beginning of 90's, an indispensable tool for the development of strategic areas, that are still emergent, but fundamental, like nanotechnology. However, this type of simulation is still of great complexity today and demands high computational power. Thus, the creation of more precise and less costly methods becomes fundamental. With the elaboration of this research, the intention is to create alternatives basis that can be used into the traditional methods of simulation of electronic structure, such as the Hartree-Fock method, GVB, among others. This essay intends to investigate part of the mathematical tools used in the calculations of electronic structure, in order to create disruptive approaches, related to the precision or velocity of the obtainment of relevant results. The new methods are based in Computational Intelligence and concepts of Functional Analysis like Wavelets.

Keywords

Nanotechnology; Computational Chemistry; Quantum Chemistry; Quantum Mechanics; Evolutionary Algorithms; Wavelets; Genetic Algorithm.

Sumário

1 INTRODUÇÃO	13
1.1. Motivação	13
1.2. Objetivos	16
1.2.1. Geral	16
1.2.2. Específicos	16
1.3. Contribuições	17
1.4. Descrição do Trabalho	18
1.5. Organização do Trabalho	19
2 FUNDAMENTOS DE MECÂNICA QUÂNTICA	21
2.1. Introdução	21
2.2. A equação de Schrödinger, a função de onda e os operadores	23
2.3. O operador Hamiltoniano, as aproximações de Born-Oppenheimer e do Campo Central	25
2.4. O Método de Hartree-Fock	27
2.5. Equações de Roothan-Hall ou Equações Seculares	34
2.6. O conceito de Funções de base em Química Computacional	35
3 INTRODUÇÃO À ANÁLISE FUNCIONAL	38
3.1. Introdução	38
3.2. Espaços de <i>Hilbert</i>	38
3.3. Operadores	42
3.4. Ortogonalização de <i>Gram-Schmidt</i>	43
3.5. Conjuntos fechados em espaços de <i>Hilbert</i>	44
3.6. Soma direta e produto tensorial	44
3.7. Espaços de Fock	46
4 INTRODUÇÃO ÀS WAVELETS	47
4.1. Introdução	47
4.2. Fundamentos	47
4.3. Análise Multirresolução	49
4.3.1. Construção das wavelets	51
4.3.2. A utilização de wavelets nessa tese	53
4.3.3. Novas funções de base propostas	55
4.3.4. Construindo funções de base	59
4.4. Polinômios de Laguerre Generalizados, uma alternativa	63
5 FUNDAMENTOS DE OTIMIZAÇÃO EVOLUCIONÁRIA	67
5.1.1. Fundamentos de programação não-linear	67
5.1.2. Modelos determinísticos e estocásticos	68
5.1.3. Fundamentos de programação evolucionária	70
5.1.4. Operadores e Parâmetros Genéticos	71
5.1.5. Mecanismo evolutivo em detalhes	72
5.1.6. Algoritmos Evolutivos com Inspiração Quântica	74

5.2. Algoritmo Evolucionário com Inspiração Quântica para Resolução de Problemas de Representação Real	74
6 FUNÇÕES DE ONDA EVOLUCIONÁRIAS E A SOLUÇÃO DA EQUAÇÃO DE SCHRÖDINGER	83
6.1. A solução da equação Schrödinger por Algoritmos Evolucionários	83
6.1.1. Breve consideração sobre o funcional de Energia de sistemas multieletrônicos	96
6.1.2. Um Estudo Bibliográfico sobre as funções de onda do estado fundamental do átomo de Hélio	96
6.1.3. O Funcional de Energia de sistemas de dois elétrons	100
6.2. Solução da equação de Schrödinger por algoritmos co-evolucionários com inspiração quântica	107
6.3. Cromossomas propostos para as evoluções que consideram utilização de funções de base	110
6.3.1. Cromossoma - para funções de base construídas a partir de polinômios de Laguerre	110
6.3.2. Cromossoma do tipo 1- para funções de base construídas a partir de wavelets	110
6.3.3. Cromossoma do tipo 2	111
7 ESTUDOS DE CASOS E DISCUSSÃO DE RESULTADOS	113
7.1. Estudo de caso 1 - Análise do desempenho de AEIQs e AEIQs Coevolutivos aplicados à construção de funções de base	114
7.2. Estudo de caso 2 -	116
7.2.1. Caso 2 a - Aplicação de AEIQs à construção de funções de base para o átomo de Lítio	116
7.2.2. Caso 2b Análise de desempenho dos Algoritmos Co-evolutivos com Inspiração Quântica	120
7.3. Estudo de caso 3	121
7.3.1. Caso 3a - Algoritmos Coevolutivos com inspiração Quântica e a ortogonalização de Gram–Schmidt	121
7.3.2. Caso 3b Algoritmos Coevolutivos com inspiração Quântica e a ortogonalização de Gram–Schmidt paramétrico	124
7.4. Caso de estudo 4	127
7.5. Caso 4a – Utilização da bases para cálculo dos átomos de Li, Be, B, C, N, O , F, Ne	127
8 CONCLUSÕES	129
9 SUGESTÕES PARA TRABALHOS FUTUROS	132
9.1.1. Redes Neurais e Neuro-evolução aplicados à solução da equação de Schrödinger	134
9.2. Cálculo Rápido das Integrais Eletrônicas por Inteligência Computacional	137
9.3. As Integrais Bi-Eletrônicas	137
9.4. Teoria do Funcional da Densidade Eletrônica (DFT)	140
9.5. Método das aproximações sucessivas aplicado	141

9.6. Possibilidade da redução de dimensionalidade em problemas de otimização	142
REFERÊNCIAS BIBLIOGRÁFICAS	144

Lista de figuras

Figura 2-1 Taxonomia dos métodos de modelagem molecular.	23
Figura 2-2 Melhor ajuste de funções Gaussianas a funções tipo Slater com diferentes parâmetros pelo critério de mínimos quadrados (SZABO & OSTLUND, 1996).	37
Figura 2-3 Aproximação de um orbital tipo Slater por combinação linear de funções Gaussianas de até três termos (SZABO & OSTLUND, 1996).	37
Figura 4-1 Decomposição do L^2 por multiresolução	51
Figura 4-2 função escala de Haar e suas integrais	60
Figura 4-3 função wavelet de Haar e suas integrais	60
Figura 4-4 Wavelet DB2, Integral da DB2 e Integral dupla da DB2 (base proposta)	62
Figura 4-5 Wavelet DB2, Integral da DB2 e Integral dupla da DB2 (base proposta)	62
Figura 4-6 Wavelet DB2, Integral da DB2 e Integral dupla da DB2 (base proposta)	63
Figura 4-7 Wavelet DB2, Integral da DB2 e Integral dupla da DB2 (base proposta)	63
Figura 5-1 Ilustra uma função que possui mínimos locais e globais.	68
Figura 5-2 Taxonomia dos métodos de otimização global, determinísticos e estocásticos.	69
Figura 5-3 Esquema de evolução dos AGs Clássicos.	71
Figura 5-4 Pseudocódigo do AEIQ-R.	75
Figura 5-5 Representação de um indivíduo composto por funções de densidade de probabilidade uniformes.	76
Figura 5-6 Funções de probabilidade acumulada associadas a cada gene de um indivíduo quântico	76
Figura 5-7 Diagrama completo do Sistema Evolutivo com Inspiração Quântica	82
Figura 6-1 Fluxograma da execução de Algoritmos Evolucionários empregados na busca por funções de onda evolucionárias.	84
Figura 6-2– O conjunto \mathcal{F}^+	86
Figura 6-3 Colaboração de vários indivíduos compondo um cromossomo	108
Figura 6-4 Esquema de co-evolução empregando três espécies	109
Figura 6-5 Cromossoma de genes reais	110
Figura 6-6 Cromossoma de genes reais	110
Figura 6-7 Cromossoma sugerido para a representação das funções de onda de cada elétron	111
Figura 7-1 Análise de convergência do AEIQ aplicado ao átomo de Li 1s22s (caso1)	115
Figura 7-2 Análise de convergência do AEIQ aplicado ao átomo de Li 1s22s (caso2)	117
Figura 7-3 a aproximação do orbital 1s do Li, e suas derivadas	118
Figura 7-4 a aproximação do orbital 2s do Li, e suas derivadas	118
Figura 7-5 Densidade eletrônica representada pelo orbital 1s	119

Figura 7-6 Densidade eletrônica representada pelo orbital 2s	119
Figura 7-7 Comparação entre os orbitais 1s e 2s do átomo de Li.	119
Figura 7-8 Comparação entre o desempenho dos algoritmos AEIQ e sua versão Coevolutiva.	121
Figura 7-9 Comparação de desempenho entre o método do estudo de caso 1 e estudo de caso 3	122
Figura 7-10 Orbitais 1s, primeira derivada e segunda derivada	122
Figura 7-11 Orbitais 2s, primeira derivada e segunda derivada	123
Figura 7-12 Comparação entre o orbitais 2s obtidos no estudo de caso 3a e 2b	123
Figura 7-13 Comparação entre as densidades radiais 1s obtidas no estudo de caso 3a e 2b	124
Figura 7-14 Comparação entre as densidades radiais 2s obtidas no estudo de caso 3a e 2b	124
Figura 7-15 Orbital 1s do Ne calculado com a base $V=1$ e $W=2$	125
Figura 7-16 Orbital 2s do Ne calculado com a base $V=1$ e $W=2$	126
Figura 7-16 Orbital 2p do Ne calculado com a base $V=1$ e $W=2$	126
Figura 9-1 Rede neural tipo feed foward.	135

Lista de tabelas

Tabela 4-1 Conjunto de coeficientes para wavelets de Daubechies	53
Tabela 4-2 Regressão linear por partes como Wavelets propostas	54
Tabela 4-3 funções de bases construídas a partir das wavelets da família Dabauchies	62
Tabela 7-1 Múltiplas configurações de algoritmos evolutivos.	116
Tabela 7-2 função de base com $V=2$ e $W=2$ para o Li 1s2s	117
Tabela 7-3 Configurações do AEIQ co-evolutivo utilizado para construir funções de base.	120
Tabela 7-4 Base para o Ne com $V=1$ e $W=2$	127
Tabela 7-5 Energia de alguns átomos calcula a partirdas funções de base com $V=1$, $W=1$	128
Tabela 7-6 Parâmetros das bases para uma aproximação com $V=1$ e $W=1$	128

1

INTRODUÇÃO

1.1.

Motivação

Diante de uma nova era de revolução, a tecnológica, a humanidade enfrenta uma quantidade de novos desafios de dimensões incomensuráveis. Isso leva à necessidade de desenvolver novos métodos científicos e tecnológicos para impulsionar o desenvolvimento de novas tecnologias. Dentro desse contexto, podem ser citados inúmeras áreas, por exemplo, a nanotecnologia, que para ser desenvolvida necessita do apoio da computação, que por sua vez necessita da nanotecnologia para evoluir.

Assim, diante da enorme quantidade de desafios presentes no dia-a-dia, surge uma área ainda emergente denominada de nanotecnologia computacional, que compreende a utilização de métodos computacionais para apoiar o desenho de sistemas nanométricos. Dentre os métodos computacionais utilizados para esse fim, destacam-se os métodos *ab-initio* utilizados como um dos pilares fundamentais da química computacional. Esses métodos são fundamentados em teorias estabelecidas na década de 30 por Edwin Schrödinger, e mais tarde aprimoradas por *Feynman*, *Dirac*, *Pauling*, *Slater* e outros físicos. Entretanto, esses métodos apenas tornaram-se passíveis de implementação computacional graças a *Hilbert*, *Von Neumann* e outros matemáticos brilhantes no decorrer de décadas.

Os métodos de simulação utilizados em Química Computacional podem beneficiar diversas áreas da ciência. Esses métodos beneficiam, por exemplo, a criação de novos fármacos, a despoluição da água e do ar, o aperfeiçoamento das tecnologias de comunicação e informação, a produção eficiente de energia, o desenvolvimento de materiais mais leves e resistentes, entre outras. No entanto, o desenvolvimento desses métodos é uma tarefa altamente complexa e custosa.

Tais métodos objetivam encontrar a solução da equação de Schrödinger, que descreve a dinâmica de um conjunto de M núcleos e N elétrons em sistemas

moleculares. No entanto, ainda não é possível obter a resolução exata dessa equação para sistemas arbitrários. Boas aproximações são conseguidas para um conjunto ainda limitado de sistemas. Uma das aproximações mais populares é conhecida como método de *Hartree-Fock*. Porém, a qualidade de tal método depende basicamente de duas aproximações: as consideradas no Hamiltoniano e as consideradas a respeito da forma funcional das funções utilizadas como tentativas (SHAVITT, 1963), (SANTOS, 1992).

Uma das implementações mais importantes no método *Hartree-Fock* para moléculas foi formalizada por *J.C. Roothaan* na década de 50, através da aproximação que ficou popularizada como “*Combinação Linear de Orbitais Atômicos (LCAO)*” (ROOTHAN, 1951). Nesse contexto, um conjunto de funções de base são combinações lineares de funções que produzem uma resposta aproximada da função de onda que representa o sistema real.

No entanto, o método de *Hartree-Fock* apresenta certas inconsistências que podem induzir a interpretações errôneas do sistema estudado, ou mesmo induzir a confusões. Uma das principais inconsistências do método de *Hartree-Fock*, segundo (NASCIMENTO *apud* COUTINHO, 2007) é a falta de simetria permutacional das funções de onda utilizadas. Tal aproximação pode não descrever bem o sistema estudado. Basicamente o método exige que para as funções serem antissimétricas com relação a permutação das suas coordenadas, as funções de onda espaciais sejam mescladas com as funções de spin, os *Spin-Orbitais* do determinante de Slater. Além disso, para sistemas moleculares é admitido no método de *Hartree-Fock* que funções de onda provenientes de Combinação Linear de Orbitais Atômicos formam funções de onda de moléculas, e isso pode não ser necessariamente verdade. Pois, uma combinação linear de orbitais atômicos pode não ser necessariamente uma autofunção do Hamiltoniano molecular, exceto no caso desses orbitais atômicos formarem uma base para o espaço de funções consideradas. Apesar disso, considerações feitas sobre as limitações do método *Hartree-Fock* guiar o desenvolvimento de métodos computacionais mais confiáveis, do ponto de vista teórico, rápidos e mais aderentes a realidade de sistemas físicos e químicos para os não é possível a utilização do *Hartree-Fock* em sua forma plena.

Métodos considerados mais precisos são baseados na inserção de variáveis que mapeiam a distância inter-eletrônica. No entanto a inserção dessas variáveis

aumenta drasticamente o esforço computacional necessário para realizar os cálculos de estrutura eletrônica. Em geral, o número de variáveis presentes numa função que considere explicitamente distâncias inter-eletrônicas é o número de elétrons somado a todas as combinações tomadas dois a dois elementos. Esse fato torna naturalmente o problema do ponto de vista computacional em um tempo finito e factível muito mais complicado.

O desenvolvimento de métodos utilizados para realizar cálculos de estrutura eletrônica é uma tarefa árida e muito complicada. Considera-se o emprego de métodos de inteligência computacional, uma vez que esses métodos vem apresentado resultados satisfatórios em milhares de aplicações, comprovadamente. Denomina-se Inteligência Computacional o conjunto de técnicas computacionais inspiradas em princípios naturais, isto é, biológicos, físicos e sociais. Essas técnicas podem ser aplicadas para resolver problemas cujas soluções sejam difíceis ou impossíveis de serem obtidas pelo uso de ferramentas tradicionais.

Neste trabalho, são aplicados alguns paradigmas de Inteligência Computacional junto a alguns preceitos de matemática aplicada, em particular, Análise Funcional. Tem-se a expectativa de que as técnicas inspiradas na natureza forneçam um novo arsenal de técnicas para o desenvolvimento de ferramentas mais adequadas à simulação *ab-initio* de sistemas polieletrônicos, isso é, técnicas que apresentem um comprometimento entre acurácia e baixo custo computacional.

Ferramentas baseadas em Inteligência Computacional têm sido aplicadas a diversos ramos da ciência. A aplicação dessas técnicas geralmente conduz a resultados que surpreendem especialistas das mais diversas áreas. Diante disso, este trabalho tem o objetivo de aplicar Inteligência Computacional como ferramenta de apoio ao cálculo da estrutura eletrônica de átomos e moléculas. Para a revisão do ferramental matemático necessário, serão discutidos alguns tópicos relacionados à Teoria dos Operadores e Teoria da Aproximação.

1.2.

Objetivos

1.2.1.

Geral

Esta pesquisa tem como objetivo principal investigar algumas propriedades do operador de Schrödinger e desenvolver procedimentos de cálculos de estrutura eletrônica, baseados em um conjunto de funções bases inédito, construído a partir de bases ortogonais de suporte compactos junto de Algoritmos Evolucionários e princípios básicos de Análise Funcional.

1.2.2.

Específicos

O objetivo desta pesquisa é analisar e desenvolver algumas novas formas de aproximar funções de onda baseadas em particularidades do operador de Schrödinger. Para tanto, será necessária a utilização de alguns teoremas fundamentais da Análise Funcional e alguns conceitos básicos, como o de espaços de Hilbert, e a noção de aproximação em espaços vetoriais de dimensão infinita, o de espaços vetoriais normados.

Deseja-se buscar diferentes formas de aproximar soluções da equação de Schrödinger, com o auxílio de Polinômios de Laguerre, Wavelets, através de procedimentos propostos neste trabalho.

Busca-se, através da revisão de alguns conceitos fundamentais da Análise Funcional e da teoria da Mecânica Quântica, apresentar um compêndio de técnicas capazes, de reunidas, convergirem para uma metodologia apta a simular, de forma mais eficaz e precisa, um sistema multieletrônico.

Além disso, é também objetivo desta pesquisa realizar a criação de ferramentas capazes de aprimorar o desempenho métodos de otimização baseados em técnicas evolutivas.

1.3.

Contribuições

Sob o ponto de vista da Inteligência Computacional, serão construídos novos modelos, tanto de otimização, quanto de aproximação, que podem eventualmente ser utilizados em outras áreas de pesquisa.

Visualizam-se também, como contribuições desta pesquisa, o estudo e integração de métodos da Matemática Aplicada com a Inteligência Computacional para a finalidade de simular sistemas polieletrônicos.

Já para a área de simulação a nanotecnologia computacional, espera-se conceber um novo método para realizar simulações mais eficientes, isto é, mais precisas e menos custosas.

Considera-se como principais resultados dessa tese:

- i. **Uma aplicação inédita do algoritmo co-evolutivo com inspiração quântica;** capaz de especificar as bases e os seus coeficientes ao mesmo tempo;
- ii. **Um novo paradigma de utilização de funções wavelets como funções de base para cálculos de estrutura eletrônica;** foi desenvolvido um arcabouço inédito que possibilita utilizar bases completas, muito simples, como a de Haar por exemplo, em cálculos de estrutura eletrônica;
- iii. **A construção de uma metodologia inédita capaz de construir funções de bases a partir de qualquer base ortogonal de suporte compacto;** foram construídas uma família inédita de bases completas a partir de wavelets para cálculos de estrutura eletrônica em geral;
- iv. **A construção de uma nova base baseada no segundo polinômio associado de Laguerre modificado;** que pode acelerar a convergência do método proposto para a solução correta. Isso pela inserção de um parâmetro não-linear na própria formação da base, variando bases possibilitando ao algoritmo realizar a variação do sub espaço no qual a solução é procurada;
- v. **Uma base cujos seus elementos podem ser obtidos por recursividade;** isso possibilita a reutilização dos elementos de base

para cálculos mais apurados; dessa forma, é vislumbrada a possibilidade da redução do esforço computacional para a aproximações de ordem superior. Apesar disso, como visto, realizar a expansão de uma função de onda em séries de Fourier pode demandar uma quantidade muito grande de integrais a serem computadas para realizar o cálculo da energia total do sistema;

- vi. **Uma metodologia para garantir que o problema de otimização é um problema restrito;** que é consequência do teorema 6-9e de seus corolários que estabelecem limites superiores e inferiores para os coeficientes das séries de Fourier;
- vii. **Um teorema que pode contribuir para a resolução de dimensionalidade de problemas de otimização em geral;**
- viii. **As discussões sobre a complexidade computacional do cálculo do range numérico do operador de Schrödinger, e os caminhos sugeridos.**

1.4.

Descrição do Trabalho

Conforme os objetivos descritos anteriormente, esta pesquisa propõe a integração de conceitos de Inteligência Computacional e Análise Funcional para criar alternativas a métodos tradicionais de simulação, presentes na Química Computacional.

Este trabalho compreende nove etapas descritas a seguir:

1. **Um estudo sobre Fundamentos de Mecânica Quântica e Química Computacional.** Para isso, realizou-se uma pesquisa bibliográfica. Nesta etapa foram estudados os princípios básicos da Mecânica Quântica aplicada a sistemas atômicos.
2. **Um estudo sobre Análise Funcional e métodos aplicados a solução da equação de Schrödinger.** Para isso, realizou-se uma pesquisa bibliográfica. Nesta etapa, foram estudadas as principais propriedades a serem consideradas neste trabalho sobre operadores e espaços normados.
3. **Um estudo sobre Algoritmos Evolutivos e suas possíveis aplicações à simulação de sistemas polieletrônicos.** Nesse estudo, além de terem sido

compiladas as informações necessárias para a implementação de Algoritmos Evolutivos, Co-evolutivos, com Inspiração Quântica (*AEIQ*) são também compreendidas as suas vantagens e desvantagens em relação aos modelos tradicionais.

4. **Um estudo teórico dos principais métodos para realizar aproximação de funções em espaços de Hilbert.** Nessa fase, foi realizado o estudo de alguns métodos que podem ser utilizados para aproximar a solução da equação de Schrödinger.
5. **Definição das alternativas estudadas para realizar a aproximação da soluções da equação de Schrödinger por Inteligência Computacional.** Nessa etapa, foi definido um conjunto de técnicas que poderiam ser utilizadas para criar e propor soluções, tais como: métodos evolucionários aplicados a parametrização de modelos baseados em Wavelets e expansões de Laguerre.
6. **Extensão do modelo de Algoritmos Evolucionários com Inspiração quântica.** Aqui se apresenta uma extensão do algoritmo evolucionário proposto em (CRUZ *et al*, 2005) para o caso co-evolucionário.
7. **Implementação do Algoritmo Evolucionário com Inspiração Quântica e de sua extensão para o caso co-evolucionário.** Nessa etapa, são implementados os algoritmos em uma linguagem de programação para que seja possível construir os sistemas propostos.
8. **Realização da construção do simulador, baseadas nas técnicas discutidas.**

1.5.

Organização do Trabalho

Esta proposta de tese está dividida em seis capítulos adicionais, descritos a seguir:

- Capítulo 2, apresenta um breve histórico da Química Computacional e Mecânica Quântica, onde são também apresentados alguns dos fundamentos necessários para a compreensão deste trabalho;
- Capítulo 3, são apresentadas, de forma detalhada, uma revisão de espaços vetoriais normados e teoria dos operadores;

- Capítulo 4, São apresentados os conceitos relativos às wavelets e polinômios de Laguerre utilizados para a criação do novo conjunto de funções de bases propostas para cálculos de estrutura eletrônica.
- Capítulo 5, são apresentadas, de forma detalhada, as principais ferramentas de otimização e técnicas utilizadas;
- Capítulo 6, discute diversos resultados considerados essenciais para realizar a integração entre os métodos propostos e Algoritmos Evolutivos;
- O capítulo 7 apresenta algumas discussões e sugestões para trabalhos futuros;
- Finalmente, o capítulo 8 apresenta as conclusões do trabalho.

2

FUNDAMENTOS DE MECÂNICA QUÂNTICA

2.1.

Introdução

Assim como em outras áreas do conhecimento humano, o uso do termo “*Química Computacional*” se popularizou devido a uma confluência de fatores científicos, históricos, tecnológicos e culturais. O desenvolvimento dessa área apenas se tornou possível devido ao poder de processamento de computadores modernos, de *softwares* sofisticados e de uma melhor compreensão de alguns princípios básicos da ciência. O potencial oferecido pela tecnologia recente de hardware e software teve como consequência o desenvolvimento de uma grande variedade de técnicas para cálculos numéricos e simbólicos. Esses métodos se estendem a muitas áreas de aplicação e aumentam o potencial das simulações, o que torna a Química Computacional um dos domínios interdisciplinares mais promissores do século XXI.

O domínio e a manipulação da matéria em escala nanométrica se mostram imprescindíveis atualmente, uma vez que, devido às suas particularidades, apresentam aplicações úteis em diversos campos de conhecimento. Para que seja possível simular o comportamento da matéria em escala atômica, são aplicados os métodos de Química Computacional baseados no formalismo da Mecânica Quântica.

Não é exagero afirmar que o ser humano tem sido envolvido por uma avalanche de consequências da Mecânica Quântica, desde seu desenvolvimento. Desde os primórdios da Física Quântica, no início do século XX, muitas formulações para representar a estrutura da matéria foram sugeridas, em parte com o objetivo de se conseguir solucionar problemas físicos em escala atômica. Uma das possíveis abordagens utilizadas para representar a estrutura da matéria é a formulação em termos da função de onda representada por Schrödinger.

Sugerida na década de 1920, ao mesmo tempo em que a Mecânica Quântica dava seus primeiros passos, a representação de Schrödinger tornou-se a abordagem mais conhecida da Física Quântica (CAETANO, 2003).

Foi em 1926 que Erwin Schrödinger e Werner Heisenberg descobriram, independentemente, os princípios básicos de um novo tipo de mecânica, que fornecia técnicas matemáticas apropriadas para lidar com a dualidade onda-partícula da matéria ou energia. A descrição de Schrödinger foi chamada de formulação Ondulatória e a de Heisenberg, Mecânica Matricial. Apesar de suas formulações bem distintas, os dois métodos são essencialmente equivalentes (DIRAC, 1929) no nível mais profundo dos conceitos físicos básicos.

Um conceito fundamental para o desenvolvimento da Mecânica Quântica foi postulado por Niels Bohr no início do século XX: a existência de estados estacionários quantizados para sistemas atômicos.

Posteriormente três cientistas construíram os alicerces da Mecânica Quântica moderna. São famosas as contribuições de Heisenberg, Schrödinger e Dirac e suas formulações são ao mesmo tempo equivalentes e complementares.

Mas, foi Erwin Schrödinger que postulou a equação responsável por descrever corretamente o movimento de sistemas quânticos. Juntando a ideia de Planck, a quantização clássica de Bohr e a ideia do comportamento ondulatório da matéria, chegou-se a uma equação admirável por conseguir descrever a dinâmica de sistemas eletrônicos.

Para que fosse possível resolver a equação de Schrödinger foi necessário o auxílio de matemáticos habilidosos e a solução somente se tornou conhecida, na década de vinte do século XX. A interpretação probabilística fornecida por Born (CAETANO, 2003), (LEACH, 2001), permitiu ao homem conhecer as “formas” geométricas do que hoje se conhece por orbitais atômicos.

A solução da equação de Schrödinger é uma função denominada função de onda, que pode-se dizer que é um ponto arbitrário em um espaço de espaço de Hilbert específico. Essa função pode ser complexa de variável real, e ainda só pode ser obtida para um conjunto restrito de casos. Para sistemas que contém mais de um elétron, um tratamento mais sofisticado é necessário.

Diversos métodos de aproximação foram propostos com o intuito de se encontrar de forma aproximada a solução da equação de Schrödinger para sistemas Polieletrônicos. Entre eles estão o método de *Hartree-Fock* e método do

Funcional da Densidade (*DFT*). Esses métodos estão em constante desenvolvimento e são aplicados rotineiramente no estudo da modelagem molecular, centro das atenções da Química Computacional. Esses métodos apenas fazem parte de um conjunto de ferramentas disponíveis para solucionar diversos problemas. São mostrados na Figura 2-1 alguns dos principais métodos utilizados em Química Computacional.

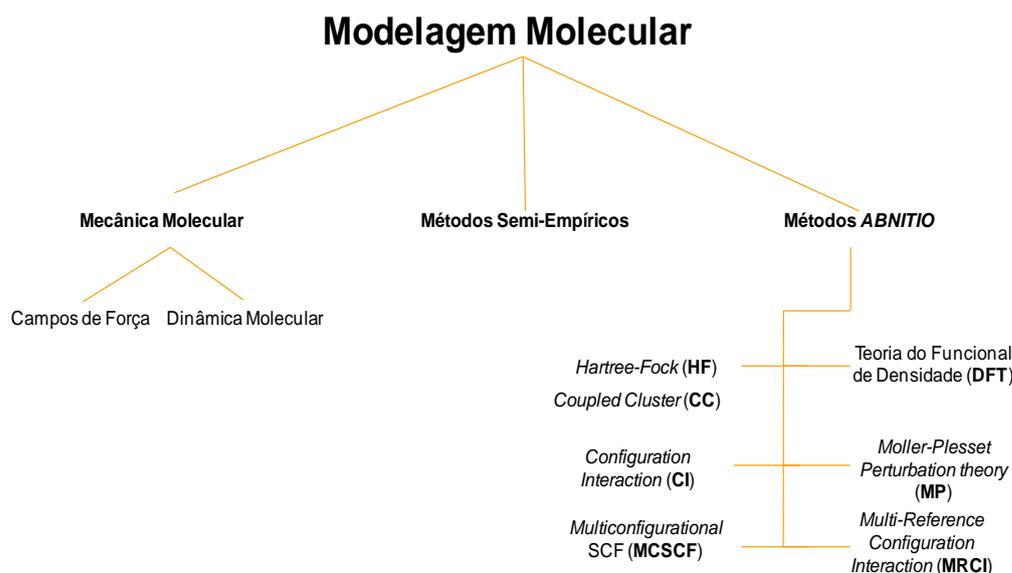


Figura 2-1 Taxonomia dos métodos de modelagem molecular.

2.2.

A equação de Schrödinger, a função de onda e os operadores

A análise de qualquer sistema eletrônico existente na natureza tem sempre como ponto inicial a equação de Schrödinger. Essa equação mescla os princípios de Hamilton com a dualidade onda-partícula.

Em Mecânica Quântica, uma partícula é associada uma função de variáveis reais com valor no campo dos complexos. Essas variáveis podem ser coordenadas espaciais e temporais.

Porém, muitas das aplicações da Mecânica Quântica direcionadas à Química Computacional, levam em conta funções de onda que não dependem do tempo. Assim, a equação de Schrödinger passa a ser uma equação de autovalor, na forma independente do tempo, como se segue:

$$H\psi = \epsilon\psi$$

ψ é denominada *função de onda*. Essa função descreve o comportamento do sistema, e pode ser complexa. E é o valor da *energia* que o sistema fornece quando o operador \mathcal{H} é aplicado sobre a função de onda que descreve o sistema, caso esse sistema se encontre num estado estacionário. \mathcal{H} é chamado de operador *Hamiltoniano* e contém termos referentes ao movimento e interação das partículas do sistema (LEACH, 2001).

A solução da equação de Schrödinger é a função de onda e toda a Mecânica Quântica moderna está baseada em tal solução (MOORE, 1972). Então, faz-se necessário definir quais são as condições para que uma função seja considerada aceitável como solução de um problema de Mecânica Quântica. A teoria apresenta os seguintes postulados:

Postulado I

Qualquer estado de um sistema dinâmico contendo N partículas pode ser descrito por uma função de onda $\psi(\mathbf{x}, t)$ onde \mathbf{x} é o vetor de coordenadas dos vários elétrons que constituem o sistema e t é o tempo.

Postulado II

A função de onda $\psi(\mathbf{x}, t)$ e sua primeira e segunda derivadas $\partial\psi(\mathbf{x}, t)/\partial\mathbf{x}$, $\partial^2\psi(\mathbf{x}, t)/\partial x^2$, devem ser contínuas e finitas.

Postulado III

Qualquer quantidade que seja fisicamente observável pode ser representada na Mecânica Quântica por um operador *hermitiano*. Um operador *hermitiano* é um operador F que satisfaz a seguinte relação:

$$\int_{-\infty}^{\infty} \psi_1(\mathbf{x})^* F \psi_2(\mathbf{x}) d\mathbf{x} = \int_{-\infty}^{\infty} \psi_2(\mathbf{x})^* F \psi_1(\mathbf{x}) d\mathbf{x} \quad 2-2$$

onde * indica o conjugado complexo. ψ_1 e ψ_2 representam os estados físicos da partícula.

Postulado V

A média ou valor esperado $\langle F \rangle$ de qualquer observável correspondente a um operador F é calculada através da fórmula

$$\langle F \rangle = \int_{-\infty}^{\infty} \psi(\mathbf{x})^* F \psi(\mathbf{x}) d\mathbf{x} \quad 2-3$$

Postulado VI

Essa formulação exige que a função de onda seja normalizada, isto é,

$$\int_{-\infty}^{\infty} \psi(x)^* \psi(x) dx = 1$$

2-4

Postulado VII

Em sistemas com condições de contorno, a função de onda deve ser nula para distâncias tendendo ao infinito, e, em sistemas sem condições de contorno, ela deve ser periódica. Dado que a função de onda satisfaz a todas as restrições anteriormente apresentadas, então o módulo do quadrado dessa função de onda define a probabilidade da referida partícula ser encontrada em uma dada região (OPPENHEIMER, 1927).

Postulado VIII

Se um operador O for capaz de permutar a posição dos núcleos ou elétrons de um sistema molecular ou atômico, sem alterar o Hamiltoniano eletrônico, então a energia do sistema é invariante frente a aplicação do operador.

2.3.**O operador Hamiltoniano , as aproximações de Born- Oppenheimer e do Campo Central**

Uma função de onda descreve o comportamento de um sistema eletrônico. No entanto, apesar da sua grande utilidade, ainda não é possível, com o ferramental teórico desenvolvido até hoje, obter a solução exata analítica da equação de Schrödinger para um sistema eletrônico qualquer. Por isso, durante décadas, surgiram vários tipos de aproximações usadas em problemas importantes, que possibilitam, ainda de forma aproximada, inferir algum tipo de comportamento de sistemas aos quais não se é possível obter soluções exatas. As principais aproximações para sistemas que possuem mais de um elétron são discutidas a seguir.

O operador Hamiltoniano que aparece na equação de Schrödinger é uma das entidades que mais contém aproximações. Muitos problemas em Química Quântica podem ser resolvidos com a equação desconsiderando os efeitos relativísticos e assumindo que a equação não depende da variável tempo. O

cálculo do Hamiltoniano, em unidades atômicas, de um sistema multieletrônico é dado pela equação (2-6) (JENSEN, 2001)

$$\mathcal{H} = - \sum_{i=1}^n \frac{1}{2} \nabla_i^2 - \sum_{A=1}^m \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^n \sum_{A=1}^m \frac{Z_A}{r_{iA}} + \sum_{i=1}^n \sum_{j>i}^m \frac{1}{r_{ij}} + \sum_{A=1}^m \sum_{B>A}^m \frac{Z_A Z_B}{r_{AB}} \quad 2-5$$

Onde, M_A é a massa do núcleo A , Z_A é o número atômico do núcleo A , r_{ij} é a distância entre os elétrons i e j , r_{iA} a distância entre o elétron i e o núcleo A ; N e M indicam respectivamente o número de elétrons e núcleos do sistema.

É possível escrever a Equação (2-6) em sua forma compacta descrita em

$$H = T_e + T_n + V_{en} + V_{ee} + V_{nn} \quad 2-6$$

onde T_e é o operador energia cinética eletrônica, T_n é o operador energia cinética nuclear, V_{en} e o operador energia potencial referente à atração elétron-núcleo, V_{ee} é o operador energia potencial referente à repulsão elétron-elétron e V_{nn} é o operador energia potencial referente à repulsão núcleo-núcleo .

A aproximação de Born-Oppenheimer, ou aproximação *adiabática*, (SZABO & OSTLUND, 1996), é formulada a partir da seguinte afirmação:

A razão entre as massas do elétron e do núcleo é suficientemente pequena de forma que os núcleos não acompanham as rápidas mudanças na trajetória dos elétrons e podem ser considerados fixos.

Assim, considera-se que os elétrons movem-se em campos com o núcleo fixo e tem-se, dentro dessa aproximação, que o termo de energia cinética nuclear T_n na Equação (2-6) é muito menor do que os outros termos e, portanto, pode ser desprezado.

Além disso, o termo de energia potencial devido à repulsão núcleo-núcleo V_{nn} é uma constante. Então, o Hamiltoniano total do sistema, assumindo-se a aproximação de Born-Oppenheimer, é dado por:

$$H = H_{ele} + V_{nn} \quad 2-7$$

onde

$$H_{ele} = T_e + V_{en} + V_{ee} \quad 2-8$$

Todas essas manipulações implicam na possibilidade de reescrever a equação $H\psi = \varepsilon\psi$, como se segue:

$$H_{ele}\psi_{ele} = \varepsilon\psi_{ele} \quad 2-9$$

Apesar da aproximação de *Born-Oppenheimer* introduzir algumas simplificações no cálculo do Hamiltoniano, ainda assim não é possível, com tal ferramenta, calcular exatamente o efeito desse operador em um sistema polieletrônico. Dessa forma, faz-se necessário introduzir outra aproximação, a aproximação de campo central.

A aproximação do campo central consiste em definir um potencial de campo médio. Dessa forma, o potencial percebido pelo *i*-ésimo elétron é resultante da distribuição média das cargas dos outros $N - 1$ elétrons, combinada a carga nuclear. Portanto, o termo de repulsão para *i* não envolve mais um somatório sobre os demais elétrons individualmente, mas sim um único termo (SANTOS, 1992), representado pela integral:

$$v(r_i) = \int \frac{\rho_{N-1}}{|r_i - r'|} dr \quad 2-10$$

Onde ρ_{N-1} representa a densidade de probabilidade de encontrar a nuvem de elétrons de $N - 1$ elétrons em r .

Através de ciclos iterativos é possível estimar ρ_{N-1} , para cada um dos elétrons presentes no sistema. Para que seja feito esse cálculo utiliza-se o método do campo auto-consistente, (SCF, do inglês *Self-Consistent Field*) (SZABO & OSTLUND, 1996).

2.4.

O Método de Hartree-Fock

A função de onda $\Psi(\mathbf{x})$ de um sistema multieletrônico pertence a um espaço de dimensão $3N$, uma vez que \mathbf{x} é o vetor de coordenadas de cada elétron no espaço. Isso mostra que é grande a complexidade computacional, e deixa claro que com as ferramentas atuais não é factível utilizar Mecânica Quântica para muitos casos. Para exemplificar essa complexidade, considere um material

contendo apenas 100 elétrons e o espaço discretizado em somente 100 pontos. Um cálculo em primeiros princípios, *ab-initio*, que envolveria a resolução de um problema de autovalores e autofunções dado pela equação (2-1), iria requerer a computação de autofunções e autovalores de uma matriz com $100^{300} \times 100^{300}$ elementos, que é um número bastante grande. Por isso é que são utilizadas aproximações analíticas das funções de onda; uma dessas aproximações é, por exemplo, o determinante de Slater.

O método de Hartree-Fock (SZABO & OSTLUND, 1996) resulta da junção da aproximação da função de onda pelo determinante de Slater, dado pela Equação (2-12), e das aproximações descritas na secção 2.3.

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = \det \begin{pmatrix} \Psi_{11} & \dots & \Psi_{1n} \\ \vdots & \ddots & \vdots \\ \Psi_{n1} & \dots & \Psi_{nn} \end{pmatrix} \quad 2-11$$

A utilização do determinante de Slater reduz uma função de onda no espaço $3N$ dimensional a um espaço de N funções no espaço tridimensional, o qual é computacionalmente tratável. Além disso, garante que a função de onda seja anti-simétrica (GOMES, 2001).

A aplicação do método de *Hartree-Fock* parte da premissa que os elétrons interagem uns com os outros ignorando a correlação eletrônica. No entanto, isso tem consequências drásticas nas interpretações das funções de onda. Em alguns casos de átomos e moléculas diatômicas, a solução numérica das equações de *Hartree-Fock* pode ser obtida por técnicas numéricas (FROOSE-FICHER, 1977). Esse tipo de solução sem aproximações (CRAMMER, 2002) conduz ao limite Hartree-Fock, e daí é possível calcular o erro proveniente da utilização de um determinante de Slater com um número finito de termos, a esse erro nomeia-se *Energia de Correlação*, e é dado pela Equação (2-13).

$$\varepsilon_{corr} = \varepsilon - \varepsilon_{HF} \quad 2-12$$

O método de *Hartree-Fock* prevê um limite superior para a energia de átomos que estejam em seu estado fundamental. Nesse contexto, o método Hartree-Fock tem uma estrutura variacional associada. Explorando essa estrutura variacional, o método de Hartree-Fock pode ser estendido para se obter uma teoria de estrutura eletrônica mais refinada, que pode ser descrita por equações de multi-configurações (FRIESECKE, 2003).

A função ψ adequada para o sistema estudado é a que minimiza a energia do sistema, dada pela Equação (2-13). A energia ε será sempre maior ou, na melhor das hipóteses igual, à energia exata ε_0 , que é a energia que seria produzida caso fosse conhecida a função de onda ψ . Tal fato pode ser enunciado pelo teorema 2.1.

Teorema 2.1 Se ψ é uma função de onda que satisfaz as condições de contorno do problema em questão, e sendo ε_0 a energia exata do estado fundamental do sistema, então:

Prova. (ver no cap 6.)

$$\frac{\int \psi \mathcal{H} \psi d\mu}{\int \psi \psi d\mu} \geq \varepsilon_0 \quad 2-13$$

onde:

ε_0 é a energia do sistema em estudo;

\mathcal{H} é o Hamiltoniano do sistema;

ψ é a função de onda em ajuste;

O valor da energia exata ε_0 é desconhecido. Porém, se forem criadas diferentes funções de onda aproximadas, $\psi_1, \psi_2, \dots, \psi_n$ para representar um dado sistema, as respostas obtidas serão as energias $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, respectivamente, e a melhor função de onda será aquela que produzir a menor energia, que certamente estará mais próxima do valor exato da energia ε_0 . Supondo-se que $\varepsilon_1 < \varepsilon_2 < \varepsilon_n$, então ε_1 estará representando mais adequadamente o sistema do que qualquer uma das outras, e $\varepsilon_2 > \varepsilon_0$.

Em linhas gerais, o método de *Hartree-Fock* baseia-se nas seguintes premissas: (a) a utilização de um determinante de Slater, dado pela equação (2-11), como função de onda aproximada; (b) a utilização do método variacional para obter uma função de onda que minimiza a energia total; e (c) a substituição das repulsões eletrônicas sofridas por cada elétron por um campo médio criado pelos outros elétrons do sistema.

A utilização da Equação (2-12) como função de onda tem uma implicação importante, pois garante que o princípio de exclusão de *Pauli* seja satisfeito (GOMES, 2001) e, portanto, que a função de onda seja anti-simétrica em relação à troca de dois elétrons. Segundo (NASCIMENTO, 2007), a antissimetria da função de onda é decorrente da simetria no Hamiltoniano e não do princípio de exclusão de *Pauli*.

Nesse modelo, cada elétron é descrito por uma função de onda Ψ denominada *spin-orbital*, dada pelo produto de uma função espacial $\chi_{11}(\mathbf{r})$ e uma função de spin

2-14

$$\Psi_{11} = \chi_{11}(\mathbf{r}) s_{11}(\eta).$$

Onde,

r é o vetor de coordenadas espaciais de um elétron e η é a coordenada de spin (alfa ou beta).

Nas funções *spins-orbitais*, os elementos do determinante de Slater têm que ser linearmente independentes, caso contrário esse determinante será zero, o que não caracteriza uma função que obedece simultaneamente aos postulados descritos na seção 2.2.

A atuação do operador Hamiltoniano eletrônico sobre a função de onda do sistema é dada pela seguinte relação:

$$\begin{aligned}
 E &= \langle \psi_a | H_\epsilon | \psi_a \rangle \\
 &= \sum_{a=1}^N \langle \psi_a | h | \psi_a \rangle \\
 &\quad + \frac{1}{2} \sum_{a=1}^N \sum_{b=1}^N (\langle \psi_a \psi_b | r_{12}^{-1} | \psi_a \psi_b \rangle - \langle \psi_a \psi_b | r_{12}^{-1} | \psi_b \psi_a \rangle)
 \end{aligned}
 \tag{2-15}$$

De acordo com o teorema 2.1, a função de onda que descreve o sistema em questão é a que minimiza o funcional dado pela Equação (2-13). No método de *Hartree-Fock*, é necessário que as funções *spin-orbitais* sejam ortogonais umas às outras e, assim, a minimização de energia deve considerar esta restrição. A condição de ortogonalidade entre as funções *spins-orbitais* é dada pela equação (2-16).

$$\langle \psi_a | \psi_b \rangle = \delta_{ab} \tag{2-16}$$

Logo,

$$\langle \psi_a | \psi_b \rangle - \delta_{ab} = 0 \tag{2-17}$$

No entanto, aqui chama-se atenção para o fato da inserção das coordenadas de spin nas funções que compõem o determinante de Slater, pois a equação de Schrödinger originalmente não leva em consideração as coordenadas de spin, e é esse um dos grandes problemas com esse método.

Tendo essa condição em vista, para se realizar o procedimento, são utilizados os multiplicadores de *Lagrange*¹. Cria-se, portanto, um funcional, Equação (2-18), que estabelece uma relação entre a energia e a restrição de ortogonalidade dada pela Equação (2-18).

¹ Técnica utilizada para resolver problemas de otimização com restrições.

$$\begin{aligned}
 F(\{\psi\}) &= E(\{\psi\}) - \sum_{a=1}^N \sum_{b=1}^N \epsilon_{ab} (\langle \psi_a | \psi_b \rangle - \delta_{ab}) \\
 &= \sum_{a=1}^N \langle \psi_a | h | \psi_a \rangle \\
 &+ \frac{1}{2} \sum_{a=1}^N \sum_{b=1}^N \left(\langle \psi_a \psi_b | r_{ab}^{-1} | \psi_a \psi_b \rangle - \langle \psi_a \psi_b | r_{ab}^{-1} | \psi_b \psi_a \rangle \right. \\
 &\left. - 2\epsilon_{ab} (\langle \psi_a | \psi_b \rangle - \delta_{ab}) \right)
 \end{aligned}
 \tag{2-18}$$

Introduzindo variações infinitesimais na função de onda, tem-se a seguinte relação:

$$\delta F(\psi) = \delta E(\psi) - \sum_{a=1}^N \sum_{b=1}^N \epsilon_{ab} (\langle \psi_a | \psi_b \rangle - \delta_{ab})
 \tag{2-19}$$

Igualando a zero, com a finalidade de satisfazer a condição de primeira ordem,

tem-se a Equação (2-22) para variações arbitrárias de $F(\psi)$ (SZABO & OSTLUND,1996)

$$\begin{aligned}
 \delta F(\{\psi\}) &= \sum_{b=1}^N \int dx \delta \psi_a^* \left\{ h \psi_a \right. \\
 &+ \sum_{b=1}^N (\langle \psi_a | r_{12}^{-1} | \psi_b \rangle - \langle \psi_b | r_{12}^{-1} | \psi_a \rangle) \psi_a \\
 &\left. - \sum_{b=1}^N \epsilon_{ab} (\langle \psi_a | \psi_b \rangle - \delta_{ab}) \right\} + \text{complexo conjugado} = 0
 \end{aligned}
 \tag{2-20}$$

A Equação (2-20) só pode ser igual a zero para variações infinitesimais se, e somente se, o termo entre chaves for igual a zero. O termo entre chaves é a equação de *Hartree-Fock*, que determina o conjunto de funções spins-orbitais que minimiza a Equação (2-20).

$$\left[\hbar + \sum_{b=1}^N (J_b - K_b) \right] \psi_a = \sum_{b=1}^N \epsilon_{ab} \psi_b, \alpha = 1, 2, \dots, N \quad 2-21$$

Os termos J_b e K_b são os operadores de Coulomb e de troca. O operador de Coulomb e de troca são mostrados nas Equações (2-22) e (2-23), respectivamente. O primeiro representa o potencial médio em um ponto devido ao elétron no orbital

ψ_b e o segundo a troca de elétrons entre os orbitais a e b .

$$J_b \psi_a = \langle \psi_b | r_{12}^{-1} | \psi_b \rangle \psi_a \quad 2-22$$

$$K_b(1) \psi_a(1) = \langle \psi_b | r_{12}^{-1} | \psi_a \rangle \psi_a \quad 2-23$$

O termo entre colchetes da Equação (2-21) é chamado de operador de *Fock*,

denotado por f . Este operador é composto pelo operador de um elétron \hbar e pelo potencial de *Hartree-Fock*, definido como

$$v^{HF} = \sum_{b=1}^N J_b - K_b \quad 2-24$$

E por fim transforma-se a Equação (2-21) na equação (2-25)

$$f \psi_a = \epsilon_a \psi_a \quad 2-25$$

Os detalhes aqui omitidos são encontrados em (JENSSEN, 1999), (SZABO & OSTLUND, 1996), bem como, os *métodos de Hartree-Fock Restrito (RHF, ou Restricted Hartree-Fock), método de Hartree-Fock Aberto e Restrito (ROHF, ou Restricted Open Hartree-Fock)*, cada um com suas particularidades e adequados a sistemas diferentes.

2.5.

Equações de Roothan-Hall ou Equações Seculares

A Equação (2-25) pode ser reescrita como a equação (2-29), que por sua vez pode ser escrita com a equação matricial (2-31). A finalidade dessa modificação é introduzir uma forma de calcular os coeficientes da combinação linear. Essa equação na forma matricial apresenta características que permitem a aplicação de técnicas numéricas eficientes para determinar os coeficientes da combinação linear e as energias dos orbitais. Popularmente, essa representação matricial é também denominada de *Equação Secular*. Os desenvolvimentos aqui omitidos podem ser encontrados em (SZABO & OSTLUND, 1996), (JENSSEN, 1999). Abaixo são ilustrados os passos para obtenção das equações seculares.

$$\psi_\alpha = \sum_n c_{in} \chi_n \quad 2-26$$

$$f \sum_n c_{in} \chi_n = \epsilon_\alpha \sum_n c_{in} \chi_n \quad 2-27$$

onde c_n são os pesos.

Esta equação pode ser rearranjada e escrita como segue:

$$\sum_n c_{in} \int \chi_m f \chi_n = \epsilon_\alpha \sum_n c_{in} \int \chi_n \chi_m \quad 2-28$$

Dessa forma,

$$\sum_n c_{in} (f_{mn} - \epsilon_\alpha S_{mn}) = 0 \quad 2-29$$

onde, f é o operador de *fock* e S é a integral de sobreposição.

E finalmente,

$$\begin{pmatrix} f_{11} - \epsilon_\alpha S_{11} & f_{12} - \epsilon_\alpha S_{12} & \dots & f_{1m} - \epsilon_\alpha S_{1n} \\ f_{21} - \epsilon_\alpha S_{21} & f_{22} - \epsilon_\alpha S_{22} & \dots & f_{2m} - \epsilon_\alpha S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{m1} - \epsilon_\alpha S_{m1} & f_{m2} - \epsilon_\alpha S_{m2} & \dots & f_{mm} - \epsilon_\alpha S_{mm} \end{pmatrix} \begin{pmatrix} C_{1k} \\ C_{2k} \\ \vdots \\ C_{mk} \end{pmatrix} = 0 \quad 2-30$$

Porém, não é aceita a solução trivial da equação secular, isto é, todos os coeficientes desconhecidos não podem ser iguais a zero. A equação secular terá solução quando o determinante, descrito pela Equação (2-31), for igual a zero.

$$\det |f_{mn} - \epsilon_\alpha S_{mn}| = 0 \quad 2-31$$

2.6.

O conceito de Funções de base em Química Computacional

Nesta seção são descritos, sucintamente, alguns aspectos relativos a funções de base. Mais detalhes podem ser encontrados nas seguintes referências: (AHLRICH & TAYLOR, 1981), (ANDZELM *et al.*, 1984), (DUNNIG & HAY, 1977), (FELLER & DAVIDSON, 1986), (FELLER & DAVIDSON, 1990), (POIRRIER *et al.*, 1986).

Os primeiros esforços no sentido de modificar as soluções numéricas das equações de *Hartree-Fock* por expressões analíticas (LÖWDIN, 1953) utilizaram funções do tipo exponencial.

Uma função de base, no contexto de química computacional, é a função resultante da combinação linear de funções linearmente independentes, que é utilizada para descrever orbitais atômicos ou moleculares. Os pesos das funções são determinados conforme discutido na seção 2.5. Dois tipos de funções de bases têm sido amplamente utilizadas, funções gaussianas e de Slater. Mas, a princípio, pode-se utilizar qualquer função de onda. Funções do tipo Exponencial, chamadas de funções de Slater, foram introduzidas na década de 30 (SLATER, 1930), e são descritas pela Equação (2-32). Tais funções propiciam uma boa representação dos orbitais atômicos, porém, apresentam algumas desvantagens em sua utilização computacional. Essas desvantagens estão relacionadas com o problema da integral de vários centros (LECHT, 2006). Resumidamente, o problema da integral de vários centros refere-se ao cálculo da interação entre diversas funções de onda.

$$\chi_i = N r^{n-1} e^{-\zeta_i r} \quad 2-32$$

As funções de Slater são semelhantes às funções orbitais do átomo de hidrogênio, sendo os polinômios de Laguerre substituídos por um único termo polinomial, r^{n-1} . Nesse caso, n refere-se ao número quântico principal (GOMES, 2001). Já N refere-se ao termo de normalização.

Nos anos 50, foram propostas em (LÖWDIN, 1953) o uso de funções gaussianas dadas pela Equação (2-33) como alternativas as funções de Slater para representar orbitais atômicos.

$$\chi_i = N x^k y^l z^m e^{-\alpha_i r^2} \quad 2-33$$

Essas funções são centradas em um núcleo A , e possuem a parte polinomial composta por três coordenadas cartesianas. A operação $k + l + m$ fornece o número quântico secundário do orbital ocupado, sendo também referido à simetria desse orbital ($s, p, d, f..$), N representa o fator de normalização. Utilizando-se essas funções pode-se contornar o problema das integrais de vários centros (LECHT, 2006), uma vez que o produto de duas gaussianas é outra gaussiana.

As funções gaussianas são as mais utilizadas atualmente para aproximação de orbitais atômicos. Uma função gaussiana não fornece uma boa representação de um orbital atômico, mas um conjunto de funções gaussianas pode representar bem. Funções gaussianas se tornaram populares por reduzir cálculos necessários na implementação de métodos da Química Quântica.

Algumas funções de base gaussianas imitam os orbitais atômicos do tipo Slater. A função de base STO-3G, por exemplo, é uma função de base que utiliza três funções gaussianas para formar cada orbital do tipo Slater. Já a função STO-4G utiliza quatro funções gaussianas para formar um orbital do tipo Slater. Alguns autores consideram impróprio chamar as funções gaussianas de orbitais atômicos do tipo gaussiano (*Gaussian Type Orbitals - GTO*), já que elas não são realmente representações exatas de orbitais. Em literaturas mais recentes (LECTH, 2006), (SANTOS, 1992), são frequentemente chamadas de primitivas gaussianas. Essas são normalmente obtidas a partir de cálculos em átomos (ex. *Hartree-Fock*). Na Figura 2-2 são mostradas as duas melhores aproximações gaussianas para uma função do tipo Slater. Além disso, na Figura 2-3 é ilustrado como uma combinação linear de funções Gaussianas de até três termos aproxima uma função tipo Slater.

A utilização de funções gaussianas apresenta algumas desvantagens, pois elas não são soluções naturais do problema e sim aproximações. As principais diferenças das Gaussianas em relação às funções de Slater são: a queda brusca da função quando r aumenta, o que dificulta a análise da região de valência e de propriedades relacionadas, uma vez que a função de onda é mal descrita nessa

região; isso prejudica a avaliação de propriedades relacionadas a região nuclear (SANTOS, 1992).

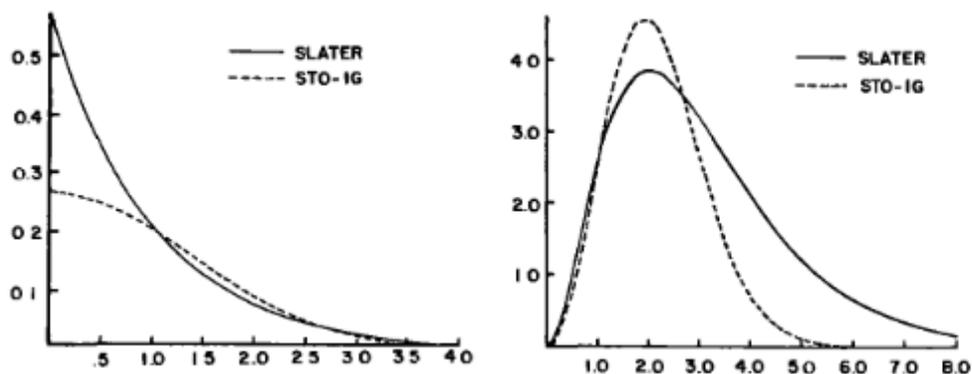


Figura 2-2 Melhor ajuste de funções Gaussianas a funções tipo Slater com diferentes parâmetros pelo critério de mínimos quadrados (SZABO & OSTLUND, 1996).

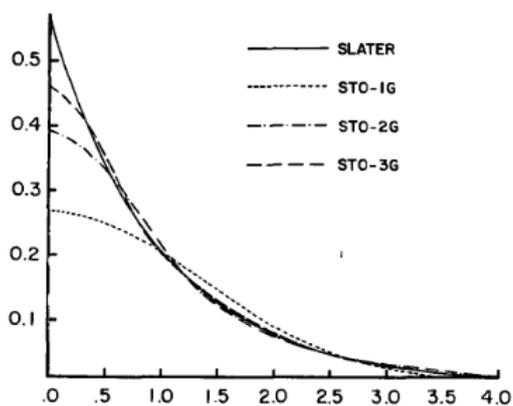


Figura 2-3 Aproximação de um orbital tipo Slater por combinação linear de funções Gaussianas de até três termos (SZABO & OSTLUND, 1996).

3

INTRODUÇÃO À ANÁLISE FUNCIONAL

3.1.

Introdução

Grande parte da Física desenvolvida no século 21 apenas pôde ser desenvolvida graças a Matemática do século 20. Em particular, com o uso dos preceitos da Análise Funcional desenvolvidos por *Hilbert* e *Von Neumman*. Nesta seção serão apresentados alguns desses conceitos, que são fundamentais para entendimento e compressão dos métodos utilizados e desenvolvidos neste trabalho. Apesar desse assunto ser de importância fundamental, todos os aspectos discutidos serão abordados de forma superficial, não mais do que o suficiente para que o trabalho seja compreendido.

3.2.

Espaços de *Hilbert*

Um espaço de Hilbert é espaço linear equipado com uma estrutura algébrica e topológica, que pode ter todos os seus elementos representados por partes simples de sua composição. Além disso, o emprego de representações baseadas em espaços de Hilbert torna possível a utilização de argumentos e conceitos geométricos, como os de ângulo e projeções, dentre outros, os que os torna, de certa forma, objetos matemáticos mais tangíveis e de fácil compreensão e assimilação.

Espaços de Hilbert desempenham um papel fundamental em toda a Física, como a mecânica quântica, por exemplo, e em várias áreas da Matemática. Historicamente sua importância foi apontada por diversos autores, mas foi especialmente *Von Neumman* quem mais destacou sua relevância.

Definição 3.0

Diz-se que um produto interno está definido num espaço vetorial \mathcal{H} se para quaisquer x, y e z em \mathcal{H} está associado a quantidade $\langle x, y \rangle$ de tal modo que a seguintes relações sejam válidas:

- i. $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$; (aditividade),
- ii. $\langle ax, y \rangle = \bar{a} \langle x, y \rangle$ para todo a ; (homogeneidade),
- iii. $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (simetria hermitiana),
- iv. $\langle x, x \rangle \geq 0$ e $\langle x, x \rangle = 0$ se, e somente se, $x = 0$ (positividade).

Definição 3.1

Um espaço de Hilbert caso seja é um espaço vetorial sobre o corpo dos complexos, munido de um produto interno, que tenha a propriedade de completude em relação a uma métrica d .

Para o caso onde o espaço vetorial \mathcal{H} é algum espaço de funções, um produto interno é pode ser dado por, sendo f e g pertencentes a \mathcal{H}

$$\langle f, g \rangle = \int f(x) \overline{g(x)} dx \quad 3-1$$

Para o caso do \mathcal{R}^n , tome $x, y \in \mathcal{R}^n$, o produto interno é dado por

$$\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i \quad 3-2$$

Exemplo de espaços de Hilbert são os espaços de dimensão finita \mathbb{C}^n , o espaço ℓ^2 , o das sequencias de quadrado somável, os espaços $L^2(M, d\mu)$, das funções quadrado integrável à *Lesbague*, espaços de *Sobolev*, espaços de *Fock*, sendo esses últimos de maior importância para o desenvolvimento desse trabalho.

Definição 3.2

Aqui define-se como uma função L^2 toda aquela que, quando definida em algum intervalo possuir integral de *Lesbague* do seu quadrado finita.

Definição 3.3

Norma de um vetor é um número real não negativo dado por

$$\|x\| = \sqrt{\langle x, x \rangle}. \quad 3-3$$

Definição 3.4

Com a definição de produto interno e norma é possível estabelecer a noção de ângulo em espaços euclidianos, dado por

$$\cos(\theta) = \frac{\langle x, y \rangle}{\|x\| \|y\|}. \quad 3-4$$

onde θ é um número real tal que $0 \leq \theta \leq \pi$.

Teorema 3.5 (Desigualdade de *Cauchy-Schwarz*)

Se x e y são dois vetores quaisquer de um espaço produto interno, então

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle. \quad 3-5$$

Definição 3.6

A distância entre dois vetores quaisquer num espaço vetorial é dado por

$$d(x, y) = \|x - y\|. \quad 3-6$$

Lema 3.7

se x e y são pontos arbitrários num espaço vetorial, então

$$\|x + y\| \leq \|x\| + \|y\|. \quad 3-7$$

Prova. Ver (Kubrusly, 2001)

Lema 3.8

Dois vetores x e y em um espaço de Hilbert são ortogonais se, e somente se,

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2. \quad 3-8$$

Prova. Ver (Kubrusly, 2001)

Teorema 3.9 (Lei do Paralelogramo)

Uma norma é associada a um produto escalar se somente se

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2. \quad 3-9$$

Prova. Ver (Kubrusly, 2001)

Teorema 3.10 (Identidade da Polarização)

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2). \quad 3-10$$

Prova. Ver (Kubrusly, 2001)

Definição 3.11

Um conjunto ortogonal, uma coleção de pontos arbitrários, x_1, x_2, \dots, x_n em um espaço vetorial se $x_i \neq 0$ para todo i , e

$$\langle x_i, x_j \rangle = 0 \quad 3-11$$

quando $i \neq j$, sendo i e j pertencentes aos naturais. Se além disso,

$$\langle x_i, x_j \rangle = 1 \quad 3-12$$

quando $i = j$, é dito que o conjunto é ortonormal.

Teorema 3.12

Todo conjunto ortogonal de pontos arbitrários num espaço vetorial \mathcal{H} é linearmente independente.

Prova. Ver (Kubrusly, 2001).

Teorema 3.13

Seja $\{e_k\}_{k \in \mathbb{N}}$ um conjunto ortonormal num espaço de Hilbert \mathcal{H} e seja x um ponto arbitrário em \mathcal{H} .

$$u_x = \sum \langle x, e_k \rangle e_k \quad 3-13$$

É um ponto único em $\text{span} \{e_k\}_{k \in \mathbb{N}}$ tal que $\|x - u_x\| \leq \|x - u\|$ para todo $u \in \text{span} \{e_k\}_{k \in \mathbb{N}}$.

Prova. Ver (Kubrusly, 2001).

Teorema 3.14 (séries de Fourier)

Seja $B = \{e_k\}_{k \in \mathbb{N}}$ um conjunto ortonormal num espaço de Hilbert \mathcal{H} . As seguintes afirmações são equivalentes:

- i. B é uma base ortonormal para \mathcal{H} ;
- ii. Todo x em \mathcal{H} tem uma única expansão em B , dada por²:

$$x = \sum \langle x, e_k \rangle e_k. \quad 3-14$$

- iii. Para todo par de pontos x, y em \mathcal{H}

$$\langle x, y \rangle = \sum \langle x, e_k \rangle \overline{\langle y, e_k \rangle}. \quad 3-15$$

- iv. Para todo x em \mathcal{H}

$$\|x\|^2 = \sum_k |\langle x, e_k \rangle|^2. \quad 3-16$$

² Essa expansão para o ponto x em \mathcal{H} é conhecida como série de Fourier de x . Os escalares dados por $\langle x, e_k \rangle$ são chamados de coeficientes de Fourier de x com relação a B .

v. Todo o subespaço linear de \mathcal{H} que inclui \mathbf{B} é denso em \mathcal{H} .

Prova. Ver (Kubrusly, 2001).

Teorema 3.15 (A representação de Riesz)

Para todo funcional linear limitado f em um espaço de Hilbert \mathcal{H} existe um único ponto y em \mathcal{H} tal que $f(x) = \langle x, y \rangle$ para todo x em \mathcal{H} . Além disso, $\|f\| = \|y\|$. Isso é chamado de representação de Riesz do funcional f em $\mathcal{B}[\mathcal{H}, \mathbb{C}]$.

Prova. Ver (Kubrusly, 2001).

3.3.

Operadores

Definição 3.16 (Operador Linear)

Seja um espaço de Hilbert H , diz-se que uma aplicação $T: H \rightarrow H$ é um operador linear se $T(\alpha f + \beta g) = \alpha T f + \beta T g$ para qualquer $\alpha, \beta \in \mathbb{C}$ para qualquer $f, g \in H$.

Definição 3.17 (Adjunto)

Seja H um espaço de Hilbert, e uma aplicação $T: H \rightarrow H$. O adjunto T^* de um operador é o único definido pela seguinte condição: $\langle f, T g \rangle = \langle T^* f, g \rangle$ para todo f e g em H .

Definição 3.18 (Auto-Adjunto ou Hermitiano)

Diz-se que um operador T é auto-adjunto se $T^* = T$.

Teorema 3.19 (Método das aproximações sucessivas)

Uma contração estrita T de um espaço métrico completo (X, d) nele mesmo tem somente um ponto fixo $x \in X$, que é o limite de toda seqüência em X da forma $\{T^n x_0\}_{n \in \mathbb{N}}$ para qualquer $x_0 \in X$.

Prova. Ver (Kubrusly, 2001).

3.4.

Ortogonalização de Gram-Schmidt

Sempre é possível transformar um conjunto de pontos linearmente independente, em um conjunto ortornormal. Para tanto, recorre-se um procedimento denominado de Ortogonalização de *Gram-Schmidt*. Esse processo é importante porque pode gerar grandes simplificações em diversos cálculos.

Teorema 3.20

Seja x_1, x_2, \dots um conjunto (finito ou infinito) de pontos linearmente independentes em um espaço de Hilbert \mathcal{H} . Então, existe um conjunto ortornormal e_1, e_2, \dots em \mathcal{H} tal que para cada inteiro n $(x_1, x_2, \dots, x_n) \sim (e_1, e_2, \dots, e_n)$. Além disso, os elementos de \mathcal{H} podem ser escritos como

$$e_1 = x_1, \quad 3-17$$

$$e_{n+1} = x_{n+1} - \alpha_1 e_1 - \dots - \alpha_n e_n, \quad 3-18$$

onde,

$$\alpha_1 = \frac{x_{n+1} \cdot e_1}{e_1 \cdot e_1}, \alpha_2 = \frac{x_{n+1} \cdot e_2}{e_2 \cdot e_2}, \dots, \alpha_n = \frac{x_{n+1} \cdot e_n}{e_n \cdot e_n} \quad 3-19$$

Prova. Ver (Kreider & Kuller, 1972).

Definição 3.21

Um conjunto ortogonal é denominado base para um espaço de Hilbert \mathcal{H} , de dimensão infinita, se todo vetor x de \mathcal{H} pode ser escrito univocamente como

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k. \quad 3-20$$

Prova. Ver (Kubrusly, 2001).

Teorema 3.22

Seja $\{g_n\}$ um sistema ortornormal no intervalo I . Então, uma função $f \in L^2$ em I , está no fecho do span de $\{g_n\}$ se e somente se:

$$f(x) = \sum_{n \in \mathbb{N}} \langle f, g_n \rangle g_n(x). \quad 3-21$$

Esses teoremas não somente são capazes de realizar enormes simplificações em muitos cálculos, quanto garantem a consistências dos métodos a serem desenvolvidos neste trabalho.

Prova. Ver (Kreider & Kuller, 1972).

3.5.

Conjuntos fechados em espaços de Hilbert**Definição 3.23**

O fecho de um subconjunto C de um espaço de Hilbert \mathcal{M} coincide com o conjunto de todos os vetores de \mathcal{M} que são pontos limites de sequências convergentes formadas por elementos de C . Pode-se entender também como fecho menor subconjunto fechado de \mathcal{M} que contém o subconjunto C .

Teorema 3.24 (Teorema do conjunto Fechado)

Um subconjunto F de um espaço de Hilbert H é fechado se, e somente se, toda sequência convergente formada por elementos de F que converge em H a um elemento x que também é elemento de F .

Prova. Ver (Kubrusly, 2001).

Definição 3.25 (Fecho do $Span$ ou $Span^-$)

Dada uma coleção de funções $\{g_n\}$, todas pertencentes a L^2 , definidas num intervalo $I \subseteq \mathbb{R}^m$, onde m é a dimensão do espaço, o $Span\{g_n\}$, pode ser entendido como o conjunto de todas as combinações lineares, de elementos de $\{g_n\}$. Denota-se

$$Span\{g_n\}^- \text{ ou } \bigvee\{g_n\} \quad 3-22$$

o menor conjunto fechado em L^2 que contém o $Span\{g_n\}$.

3.6.

Soma direta e produto tensorial

Dados dois espaços de Hilbert \mathcal{H}_1 e \mathcal{H}_2 , define-se soma direta ou soma ortogonal $\mathcal{H}_1 \oplus \mathcal{H}_2$ como o conjunto de pares ψ_1 e ψ_2 pertencente a $\mathcal{H}_1 \times \mathcal{H}_2$ junto com o seu produto escalar

$$\langle (\varphi_1, \varphi_2), (\psi_1, \psi_2) \rangle = \langle \varphi_1, \psi_1 \rangle + \langle \varphi_2, \psi_2 \rangle. \quad 3-23$$

É possível mostrar que $\mathcal{H}_1 \oplus \mathcal{H}_2$ é também espaço de Hilbert. Além disso, \mathcal{H}_1 pode ser identificado como $\{(\psi_1, 0) \mid \psi_1 \in \mathcal{H}_1\}$ e é possível assegurar

que \mathcal{H}_1 como subespaço de $\mathcal{H}_1 \oplus \mathcal{H}_2$, da mesma forma que para \mathcal{H}_2 . Também é usual escrever $\psi_1 \oplus \psi_2$ ao invés de (ψ_1, ψ_2) . Ver (Kubrusly, 2001).

De forma mais geral, seja $\mathcal{H}_j, j \in \mathbb{N}$, se uma coleção contável de espaços de *Hilbert*, define-se

$$\bigoplus_j^\infty \mathcal{H}_j = \left\{ \sum_j \psi_j \mid \psi_j \in \mathcal{H}_j, \sum_j \|\psi_j\|^2 < \infty \right\}, \quad 3-24$$

que é um espaço de Hilbert com o seguinte produto interno

$$\left\langle \sum_i \varphi_i, \sum_j \psi_j \right\rangle = \sum_i \langle \varphi_i, \psi_i \rangle. \quad 3-25$$

Ver (Teschl, 2009).

Da mesma forma, se \mathcal{H}_1 e \mathcal{H}_2 são dois espaços de *Hilbert*, define-se o produto tensorial como:

Os produtos da forma $\psi_1 \otimes \psi_2$ de elementos ψ_1 em \mathcal{H}_1 e ψ_2 em \mathcal{H}_2 , que são classes de equivalências para (ψ_1, ψ_2) , (Teschl, 2009). para uma construção do produto tensorial de espaços de Hilbert arbitrários ver (Kubrusly, 2006).

Sejam $f: \mathcal{R}^n \rightarrow \mathbb{C}$ e $g: \mathcal{R}^m \rightarrow \mathbb{C}$ duas funções em L^2 . É usual definir o produto tensorial de f por g , designado por $f \otimes g$, como sendo a função $z: \mathcal{R}^{n+m} \rightarrow \mathbb{C}$ definida por

$$f \otimes g = f(x)g(y) \quad 3-26$$

onde x e y são pontos arbitrários de \mathcal{R}^n e \mathcal{R}^m respectivamente.

Define-se ξ um ponto \mathcal{R}^{n+m} , $\xi = (x_1 \dots x_n, y_1 \dots y_m)$ que pode ser identificando, no produto cartesiano $\mathcal{R}^n \times \mathcal{R}^m$, e naturalmente, escreve-se $\xi = (x, y)$.

A aplicação de cada par de $f: \mathcal{R}^n \rightarrow \mathbb{C}$ e $g: \mathcal{R}^m \rightarrow \mathbb{C}$, associada a função $f \otimes g: \mathcal{R}^{n+m} \rightarrow \mathbb{C}$, é bilinear. Prova ver (Kubrusly, 2006).

Definição 3.26

Uma aplicação $f: V \times V \rightarrow \mathcal{R}$ é dita bilinear se para qualquer x, y e z em V em \mathcal{R} se:

- i. $f(x + y, z) = f(x, z) + f(y, z)$;
- ii. $f(z, x + y) = f(z, x) + f(z, y)$;

$$\text{iii. } f(kx, y) = kf(x, y) = f(x, ky).$$

Teorema 3.27

Se f e g pertencem espaços de Hilbert \mathcal{H}_1 e \mathcal{H}_2 de funções do \mathcal{R}^n e \mathcal{R}^m , respectivamente, então $f \otimes g$ está em $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Prova. Ver (Teschl, 2009).

Teorema 3.28

Se B_1 e B_2 são bases dos espaços de Hilbert \mathcal{H}_1 e \mathcal{H}_2 de funções do \mathcal{R}^n e \mathcal{R}^m , então $B_1 \otimes B_2$ é base de $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Prova. Ver (Kubrusly, 2006).

3.7.**Espaços de Fock**

Um espaço de *Fock* é um espaço de *Hilbert* composto de uma soma direta de produtos tensoriais de espaços de *Hilbert*

$$\mathcal{F} = \bigoplus_{n=0}^{\infty} S\mathcal{H}^{\otimes n},$$

onde n é um número natural;

S é um operador que simetriza ou antissimetriza um produto de n elementos quaisquer.

4

INTRODUÇÃO ÀS WAVELETS

4.1.

Introdução

Neste capítulo será apresentada, com pouco rigor, parte da teoria utilizada para realizar a construção das novas funções de base propostas a partir das wavelets (DAUBECHIES, 1992).

Um ponto que merece destaque é que as séries wavelets são apenas um caso particular de séries de Fourier, onde a base é um conjunto composto de funções wavelets. A utilização de séries de Fourier possui diversas aplicações em todos os ramos da ciência. No entanto, por razões históricas e outras práticas é utilizada apenas um tipo de série de Fourier, a de senos e cossenos. Séries de Fourier são ferramentas muito mais gerais do que aquelas apresentadas compostas por senos e cossenos. Muitas séries de Fourier podem ser construídas com a utilização de funções que não possuem sequer descrição por funções primitivas, nesse contexto, que possam ser escritas como, seno, cosseno, exponenciais, potências, e tantas outras.

4.2.

Fundamentos

Uma questão que surge naturalmente quando se refere às expansões em wavelets é: por que utilizar wavelets como base ao invés de senos e cossenos, ou outras?

A resposta para essa questão depende basicamente do domínio da função a ser aproximada. Pois, a teoria desenvolvida no capítulo 3 nos fornece um ferramental para expandir funções que são quadraticamente integráveis. Nesse ponto, é importante mencionar que ser ou não integrável depende, entre outras coisas, do domínio de integração. Nesse contexto, afim de ilustrar essa afirmação considere a função, $f: \mathcal{R} \rightarrow \mathcal{R}$, onde $f(x) = \text{sen}(x)$, é imediato verificar que a

integral $\int_0^{\infty} |\text{sen}(x)|^2 dx$ não está definida. Logo, a função $\text{sen}(x)$ não pertence ao conjunto $L^2(0, \infty)$ e logo não pode ser utilizada em nenhuma circunstância como base para esse conjunto.

As expansões em wavelets são séries de Fourier compostas por funções Wavelets. Funções wavelets são bases ortogonais que conseguem gerar o $L^2(-\infty, \infty)$ sobre a operação dos operadores *Shift*, D , e *Translação*, T , equações (4-1) e (4-2). Esses operadores são operadores lineares definidos em $L^2(-\infty, \infty)$.

Sejam as funções f, g, z, ψ e $\psi_{m,n}$ definidas em $L^2(-\infty, \infty)$. Define-se os operadores *Shift* e *Translação*, (LEVAN & KUBRUSLY, 2006), (DAUBECHIES, 1992) da seguinte forma

$$Df(x) = g(x) \tag{4-1}$$

$$\text{onde, } g(x) = \sqrt{2} f(2x)$$

$$Tf(x) = z(x) \tag{4-2}$$

$$\text{onde, } z(x) = f(x - 1)$$

$$\begin{aligned} \psi_{m,n}(x) &= \sqrt{2}^m \psi(2^m x - n) \\ &= D^m T^n \psi(x) \end{aligned} \tag{4-3}$$

Em *Abstract wavelets generated by Hilbert space shift operators* (LEVAN & KUBRUSLY, 2006), é mostrado que o conjunto $\{D^m T^n \psi(x)\}_{(m,n) \in \mathbb{Z}}$ é base em $L^2(-\infty, \infty)$, desde que ψ seja uma função wavelet.

A propriedade mais notável das wavelets é que tais funções são ortogonais às suas translações e dilatações. Assim, todas as funções pertencentes a $L^2(-\infty, \infty)$ podem ser aproximadas, com precisão arbitrariamente alta, a partir de combinações lineares das *wavelets* $\psi_{m,n}$. Logo, $\forall f \in L^2(-\infty, \infty)$ existe uma aproximação que pode ser escrita como,

$$f(x) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \alpha_{m,n} \psi_{m,n} \tag{4-4}$$

onde $\alpha_{m,n}$ é dado por:

$$\alpha_{m,n} = \int_{-\infty}^{\infty} f(x) \psi_{m,n}(x) dx \quad 4-5$$

As wavelets, que ainda não possuíam esse nome, foram introduzidas em 1910 pelo matemático Alfred Haar, na sua tese de doutorado sobre a teoria do sistema de funções ortogonais (HAAR, 1910). Em ciências naturais, é muito comum a ideia de decompor um problema em suas partes mais simples, estudá-las individualmente, e então construir um modelo do problema a partir do entendimento de cada uma dessas componentes, e é essa a ideia por trás da utilização de bases para decompor espaços de Hilbert de funções. Graças às bases wavelets também é possível a introdução de uma técnica chamada análise multirresolução.

4.3.

Análise Multirresolução

Um *framework* natural para estudar e entender as wavelets é a análise multirresolução (*Multiresolution Analysis - MRA*), (MALLAT, 1989) (DAUBECHIES, 1992) (PAGAMISSE, 2004), (STRANG, 1989) que é basicamente uma construção matemática que caracteriza as wavelets de forma geral. A partir desse *framework* é possível decompor um espaço de Hilbert em vários subespaços de forma sistemática. O objetivo da análise multirresolução é entender como representar uma função f em $L^2(-\infty, \infty)$ a partir de bases para cada subespaço de um espaço de Hilbert. Essa análise é construída a partir dos seguintes axiomas:

$$\{0\} \subset \dots \subset \{V_{-1}\} \subset \{V_0\} \subset \{V_1\} \subset \dots \subset L^2(-\infty, \infty) \quad (i)$$

$$\overline{\bigcup_{j=-\infty}^{\infty} V_j} = L^2(-\infty, \infty) \quad (ii)$$

$$\bigcap_{j=-\infty}^{\infty} V_j = \{0\} \quad (\text{iii})$$

$$\{\varphi(x-k)\}_{k \in \mathbb{Z}} = L^2(-\infty, \infty) \text{ é uma base ortonormal para } V_0 \quad (\text{iv})$$

$$f \in V_j \Leftrightarrow f(2 \cdot) \in V_{j+1} \quad (\text{v})$$

Esse conjunto de axiomas descreve uma sequência de subespaços V_j em $L^2(\mathcal{R})$ tal que o fecho da união é igual a $L^2(\mathcal{R})$. A partir dessa sequência são construídos os espaços W_j definidos como o complemento ortogonal de V_j em V_{j+1} , isso é $V_j \perp W_j$ e

$$V_{j+1} = V_j \oplus W_j, \quad 4-6$$

$$W_j = V_{j+1} \ominus V_j \quad 4-7$$

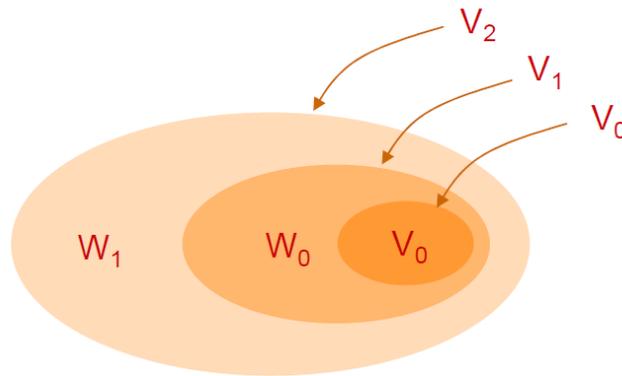
Considerando dois espaços V_{j_0} e V_j , onde $j > j_0$ aplicando a equação 4-5 recursivamente encontra-se:

$$V_j = V_{j_0} \oplus \left(\bigoplus_{j=j_0}^{j-1} W_j \right) \quad 4-8$$

Além disso, a partir de (4-7) deduz-se que,

$$L^2(\mathcal{R}) = \bigoplus_{j=-\infty}^{\infty} W_j \quad 4-9$$

Na Figura 4-1 é mostrado um esquema de representação de gráfica do esquema de multiresolução proposto pelo conjunto de axiomas i), ii), iii), iv), v) e pela equação (4-7).

Figura 4-1 Decomposição do L^2 por multiresolução

4.3.1.

Construção das wavelets

São diversas as formas e técnicas de se realizar a construção de funções wavelets. Essas técnicas podem ser encontradas em (MALLAT, 1989), (DAUBECHIES, 1992), (WEISS, 1996). Geralmente, as wavelets que são ortogonais são originadas da solução da chamada equação de dilatação (STRANG, 1989) definida na equação (4-14):

onde, D e T são os operadores de Dilatação e Translação, respectivamente.

Para ilustrar a utilização das equações de dilatação, vamos definir:

$$V_0 = \overline{\text{span}\{T^k g\}: k \in \mathbb{Z}} \quad 4-10$$

$$= \overline{\text{span}\{g(\cdot - k)\}: k \in \mathbb{Z}} \quad 4-11$$

$$V_1 = \overline{\text{span}\{T^k Dg\}: k \in \mathbb{Z}} \quad 4-12$$

$$= \overline{\text{span}\{g(2 \cdot - k)\}: k \in \mathbb{Z}} \quad 4-13$$

Daí, se uma dada função $g \in \{V_0\}$, $g \in \{V_0\} \subset \{V_1\}$, então,

$$g(x) = \sum_{k=0}^n c_k g(2x - k). \quad 4-14$$

Essa equação funcional é dita equação de dilatação, que pode ser formulada como uma equação de autovalor. Imposições sobre o conjunto de coeficientes determina a família e forma das wavelets.

Nesta tese, soluções das equações de dilatação são obtidas da seguinte forma: sejam os operadores V e W definidos em (4-15) e (4-16)

$$(V\varphi_0)(x) = \sum_{k=0}^n c_k \varphi_0(2x - k) \quad 4-15$$

Onde φ_0 é uma base para V_0 .

$$(W\varphi_0)(x) = \sum_{k=0}^n c_k (-1)^k \varphi_0(2x - k) \quad 4-16$$

As soluções da equação de dilatação (4-15) e (4-16): são pontos fixos dos operadores V e W . A solução de cada uma dessas equações é o limite da sequência de potências de cada um dos operadores em questão. Logo,

$$\varphi(x) = \lim_{n \rightarrow \infty} V^n \varphi_0(x) \quad 4-17$$

$$\psi(x) = \lim_{n \rightarrow \infty} W^n \varphi_0(x) \quad 4-18$$

As funções φ e ψ são respectivamente as funções escala e wavelets.

Nesta tese, $\varphi_0(x)$ é definida como a função de Haar. Esse procedimento muitas vezes é referido como *cascade algorithm* (BURRUS, 1988), (MALLAT, 1989). Será utilizados neste trabalho o conjunto de coeficientes $C = \{c_0 \dots c_n\}$, que caracteriza a família Dabuchies (DABUCHIES, 1992), dados pela

Tabela 4-1. Nessa tabela, os coeficientes para as wavelets Daubechies estão separados em cada coluna:

Tabela 4-1 Conjunto de coeficientes para wavelets de Daubechies

D2	D4	D6	D8	D10	D12	D14	D16	D18	D20
1	0,683013	0,470467	0,325803	0,226419	0,157742	0,110099	0,076956	0,05385	0,037717
1	1,183013	1,141117	1,010946	0,853944	0,699504	0,560791	0,442467	0,344834	0,266122
	0,316987	0,650365	0,892201	1,024327	1,062264	1,031148	0,955486	0,855349	0,745575
	-0,18301	-0,19093	-0,03958	0,195767	0,445831	0,664372	0,827817	0,929546	0,973628
		-0,12083	-0,26451	-0,34266	-0,31999	-0,20351	-0,02239	0,18837	0,397638
		0,049818	0,043616	-0,0456	-0,18352	-0,31684	-0,40166	-0,41475	-0,35334
			0,046504	0,109703	0,137888	0,100847	6,68E-04	-0,13695	-0,27711
			-0,01499	-0,00883	0,038923	0,114003	0,182076	0,210068	0,180127
				-0,01779	-0,04466	-0,05378	-0,02456	0,043453	0,131603
				4,72E-03	7,83E-04	-0,02344	-0,06235	-0,09565	-0,10097
					6,76E-03	0,01775	0,019772	3,55E-04	-0,04166
					-1,52E-03	6,08E-04	0,012369	0,031624	0,04697
						-2,55E-03	-6,89E-03	-6,68E-03	5,10E-03
						5,00E-04	-5,54E-04	-6,05E-03	-0,01518
							9,55E-04	2,61E-03	1,97E-03
							-1,66E-04	3,26E-04	2,82E-03
								-3,56E-04	-9,70E-04
								5,56E-05	-1,65E-04
									1,32E-04
									-1,88E-05

4.3.2.

A utilização de wavelets nessa tese

A utilização de funções limite da sequência de potências definidas na equações (4-17) e (4-18) levam a obtenção de wavelets que não podem ser descritas com o auxílio de funções primitivas, isto é, não possuem descrições em termos de funções como senos, cossenos, exponenciais, logaritmos, etc. Porém, esse procedimento é capaz de fornecer a imagem de uma wavelet num conjunto de pontos discretos. Mas, é impraticável realizar o cômputo desse conjunto de pontos de imagem toda vez que seja necessário a utilização da wavelet. Assim, nessa tese, foi introduzido um artifício para realizar e minimizar o esforço computacional para a utilização das funções wavelets.

O artifício utilizado é fundamentado na garantia fornecida em que o procedimento sugerido dá origem a funções contínuas, de forma mais clara, os pontos fixos dos operadores definidos em V e W em (4-17) e (4-18) convergem

para funções contínuas. Para que seja construída uma função com o comportamento semelhante ao de uma wavelet em termos computacionais nessa tese sugerimos obter a imagem da wavelet em um conjunto de pontos igualmente espaçados e realizar uma interpolação linear por partes. Esse procedimento pode ser de grande valia, uma vez que, do ponto de vista computacional as funções wavelets agora possuem uma forma analítica fácil de tratar e manipular. O argumento de continuidade é essencial para garantir a consistência dessa representação. Pois, uma vez que a função a ser representada é uma função contínua, é notável e evidente a proximidade entre os pontos na vizinhança da sua imagem, e isso sugere a razoabilidade da utilização de uma interpolação linear por partes, definida pela equação (4-18). Assim, tanto as funções escala quanto as wavelets passam a serem descritas como:

$$\psi(x) = \begin{cases} a_0x + b_0, & \text{se } x_0 \leq x < x_1 \\ a_1x + b_1, & \text{se } x_1 \leq x < x_2 \\ \vdots & \\ a_nx + b_n, & \text{se } x_n \leq x < x_{n+1} \end{cases} \quad 4-19$$

Onde,

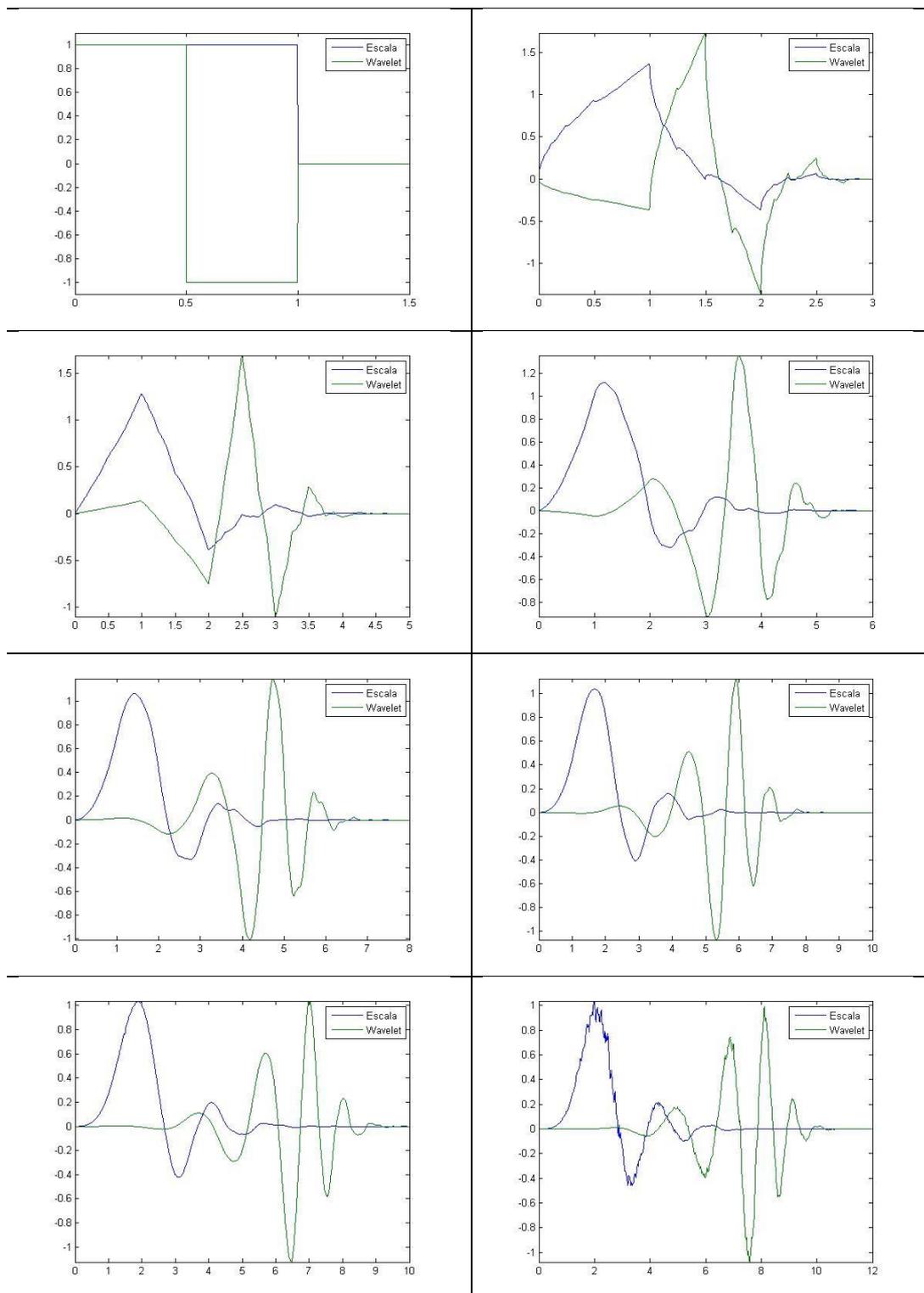
a_n e b_n representam os coeficientes da regressão linear por partes.

A forma das wavelets e funções escala definidas pela equação (4-17) é também uma das contribuições dessa tese, apesar de ser algo relativamente simples. Esse tratamento tem a vantagem de permitir o tratamento analítico e o cômputo de propriedades, tais como integrais e derivadas necessárias durante o processo de cálculo das propriedades dos sistemas quânticos estudados. De outra maneira, do ponto de vista computacional, o uso de muitas famílias de funções wavelets não seria viável, isso devido ao esforço computacional empreendido para realizar os cálculos de limites de sequência.

Nas figuras da Tabela 4-2 são mostrados os resultados da aplicação do procedimento proposto nessa tese para o cálculo de cada wavelet e função escala construída a partir da utilização dos coeficientes da

Tabela 4-1 e procedimento proposto:

Tabela 4-2 Regressão linear por partes como Wavelets propostas



4.3.3.

Novas funções de base propostas

Apesar de todo o desenvolvimento da teoria de wavelets e do esforço para representar os elementos do espaço de Hilbert $L^2(-\infty, \infty)$, com a multiresolução, por exemplo, pode não ser viável a utilização do método variacional em conjunto com wavelets. A justificativa para essa afirmação baseia-se no fato dos operadores diferenciais serem não limitados, como o laplaciano presente no Hamiltoniano.

No entanto, nesse capítulo é proposto um caminho para viabilizar a utilização dessa teoria em cálculos quânticos. As funções wavelets são modificadas dando origem a uma nova classe de funções.

A nova classe de funções de base parte diretamente das restrições que impõe:

1. A dupla diferenciabilidade da função de onda, isto é, a segunda derivada da função de onda deve existir em todos os pontos do domínio. Nos casos trabalhados esse domínio é a semi-reta positiva;
2. A convergência assintótica da imagem da função de onda para 0.

O arcabouço desenvolvido pressupõe que as derivadas de ordem 2 das funções de onda sejam aproximadas, e não as funções que seriam soluções da equação de Schrödinger, isso é a função de onda. Admitindo que existe uma série de wavelets capaz de aproximar as derivadas segundas da função de onda, basta integrar a série obtida para realizar a aproximação derivada da função e integrar novamente para obter a série de funções que seria uma aproximação para a solução desejada. Em poucas palavras, $\Psi \in \overline{C^2(0, \infty) \cap L^2(0, \infty)}$.

A não viabilidade da utilização direta das funções wavelets, ainda que fossem obtidas facilmente a baixo custo computacional, tal como já mencionado, resulta das funções wavelets não possuírem necessariamente, derivadas em todos os pontos do domínio, como exemplo, cita-se a função de Haar, que é descontínua, ou mesmo a DB4, contínua, porém com derivadas de ordem superior descontínuas.

Por isso, opta-se pela utilização de wavelets duplamente integradas como funções de base. Infelizmente, não é possível garantir que as funções wavelets duplamente integradas sejam bases ortogonais. Isso traz perdas irreparáveis, como a impossibilidade da utilização do teorema de Fourier para a especificação dos coeficientes da série, dentre outras. Devido ao princípio variacional esse problema torna-se contornável.

A proposição 4-1 é a base para a construção das funções de base propostas nesta tese. Essa proposição ilustra que se uma função é definida $L^2(0, \infty)$ e tem derivadas segundas também em $L^2(0, \infty)$ é possível construir uma nova base a partir de outra base já existente.

Proposição 4-1 (A construção de uma base nova) .

Sejam $\psi, g \in L^2(0, \infty)$, $g(x) = \frac{\partial^2}{\partial x^2} \psi(x)$ e Φ_k um elemento de uma base ortonormal para $L^2(0, \infty)$. Existe uma sequência $\{\alpha_k\}_k^\infty$ e um elemento de uma base de Riesz $\overline{\Phi}_k(x, K_0, K_1) = \int_0^x \int_0^{x'} \Phi_k(x'') dx'' dx' + K_0 x + K_1$ que aproxima algum ψ com segunda derivada igual g . Onde $K_0, K_1, \in \mathfrak{R}$.

Prova.

Pelo teorema fundamental do cálculo, se uma dada função $f: \mathcal{R} \rightarrow \mathcal{R}$ é integrável em $[0, x]$, $x \in \mathfrak{R}$ seguem as seguintes relações:

$$\int_0^x f(x) dx = F(x) - F(0),$$

$$F(x) = \int_0^x f(x) dx + F(0),$$

$$\int F(x) dx = G(x)$$

$$\int_0^x F(x) dx = \int_0^x \int_0^{x'} f(x'') dx'' dx' + F(0)x + G(0),$$

Logo,

$$\frac{\partial^2}{\partial x^2} \psi(x) = \lim_{n \rightarrow \infty} \sum_k^n \alpha_k \Phi_k(x),$$

$$\frac{\partial}{\partial x} \psi(x') dx' = \int_0^x \frac{\partial^2}{\partial x^2} \psi(x') dx' + K_0,$$

$$= \lim_{n \rightarrow \infty} \int_0^x \sum_k^n \alpha_k \Phi_k(x) dx' + K_0$$

$$= \lim_{n \rightarrow \infty} \sum_k^n \alpha_k \int_0^x \Phi_k(x) dx' + K_0,$$

$$\int \int \frac{\partial^2}{\partial x^2} \psi(x'') dx'' dx' = \lim_{n \rightarrow \infty} \sum_k^n \alpha_k \int_0^x \int_0^{x'} \Phi_k(x'') dx'' dx' + K_0 x + K_1,$$

$$= \lim_{n \rightarrow \infty} \sum_k^n \alpha_k \int_0^x \int_0^{x'} \Phi_k(x'') dx'' dx' + K_0 x + K_1,$$

$$\psi(x) = \lim_{n \rightarrow \infty} \sum_k^n \alpha_k \int_0^x \int_0^{x'} \Phi_k(x'') dx'' dx' + K_0 x + K_1,$$

Daí,

$$\overline{\Phi}_k(x, K_0, K_1) = \int_0^x \int_0^{x'} \Phi_k(x'') dx'' dx' + K_0 x + K_1.$$

$$\psi(x) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \alpha_k \overline{\Phi}_k(x), \blacksquare$$

A descrição das wavelets como funções lineares por partes foi artifício utilizado para tornar possível a estimativa das integrais das wavelets em pontos específicos. Apesar de existirem algoritmos capazes de estimar a imagem das wavelets, até o momento não está disponível um método consistente para realizar a estimação direta da imagem da aplicação de operadores nas wavelets, ainda que sejam wavelets mais simples, como a de Haar, por exemplo.

De acordo com (DABUCHIES, 1988), (DABUCHIES, 1992) as funções wavelets originadas do procedimento descrito na seção 4.3.1 são contínuas e de suporte compacto. Por suporte compacto entende-se que existe uma região fechada e limitada no domínio de uma dada função definida nos reais para o qual a partir de uma certo ponto a imagem da função seja igual a zero.

Uma das condições fundamentais de admissibilidade para uma função de onda é que seja possível integrar o seu quadrado, ou seja, $\int_0^\infty f(x)^2 dx < \infty$, isso implica que $\lim_{x \rightarrow \infty} f(x) = 0$ e por outro lado se é utilizado $\overline{\Phi}_k(x)$ como elemento de base o $\lim_{x \rightarrow \infty} \overline{\Phi}_k(x, K_0, K_1) = 0$ para todo k , e isso pode ser feito, com especificação correta das constantes K_0, K_1 .

Proposição 4.2

Se Φ_k é uma função de suporte compacto então existem constantes K_0, K_1 que garantem que $\lim_{x \rightarrow \infty} \overline{\Phi}_k(x, K_0, K_1) = 0$, onde

$$\overline{\Phi}_k(x, K_0, K_1) = \int_0^x \int_0^{x'} \Phi_k(x'') dx'' dx' + K_0 x + K_1.$$

Prova.

Proposição 4.3

Existem constantes K_0, K_1 que fazem com que

$$\bigvee_k \overline{\Phi_k}(x, K_0, K_1) = (\overline{C^2(0, \infty) \cap L^2(0, \infty)}, \| \cdot \|_2).$$

4.3.4.

Construindo funções de base

Nesta seção serão construídas as funções de base $\overline{\Phi_k}(x, K_0, K_1)$, iniciando pela wavelet mais simples possível, a de Haar, a partir da proposição 4.2. Posteriormente o procedimento será generalizado para a construção de funções de base criadas a partir da família Daubechies, ou outras de suporte compacto existentes na literatura.

A função de Haar é definida pela equação 4-19. Para realizar translações e contrações mais gerais das wavelets de Haar é utilizada a equação 4-20:

$$\phi(x) = \begin{cases} 0, & \text{se } -\infty \leq x < 0 \\ 1, & \text{se } 0 \leq x < 1 \\ 0, & \text{se } 1 \leq x < \infty \end{cases} \quad (4-20)$$

$$\phi(\alpha x + \beta) = \begin{cases} 0, & \text{se } -\infty \leq x < -\beta/\alpha \\ 1, & \text{se } -\frac{\beta}{\alpha} \leq x < \frac{1-\beta}{\alpha} \\ 0, & \text{se } \frac{1-\beta}{\alpha} \leq x < \infty \end{cases} \quad (4-21)$$

onde, α e β são fatores de dilatação e translação, respectivamente.

Para realizar a aplicação da proposição 4-1 e construir a base $\overline{\Phi_k}(x, K_0, K_1) = \int_0^x \int_0^{x'} \phi_k(x'') dx'' dx' + K_0 x + K_1$ é necessário realizar sempre o cômputo das seguintes quantidades:

$$\int_0^x \phi(\alpha x' + \beta) dx' + K_0 \quad \int_0^x \int_0^{x'} \phi(\alpha x'' + \beta) dx'' dx' + K_0 x + K_1 \quad K_0 \quad K_1$$

Nas figuras Figura 4-2 e Figura 4-3 representadas pelas curvas azuis, verdes e vermelhas estão as wavelets, as integrais das wavelets e as integrais duplas, respectivamente, tanto das funções escala quanto das wavelets. As curvas vermelhas nas figuras da Tabela 4-3 são as bases construídas a partir da aplicação da proposição 4-1.

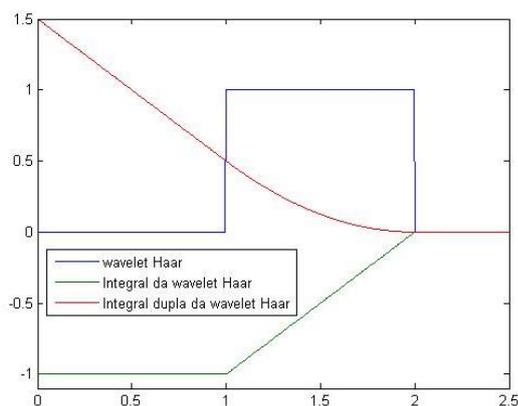


Figura 4-2 função escala de Haar e suas integrais

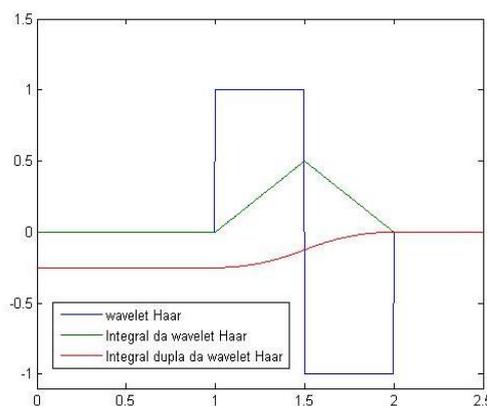


Figura 4-3 função wavelet de Haar e suas integrais

Para as wavelets descritas como funções lineares por partes são utilizados alguns artifícios analíticos como o uso da função indicador, por exemplo, $U_g(x')$. Notar que a_k e b_k são os coeficientes da regressão linear por partes correspondentes a cada uma das wavelets de suporte compacto representadas.

Assim,

$$U_g(x') = \begin{cases} 1 & ,se\ 0 \leq x' \leq g \\ 0 & ,\ caso\ contrário \end{cases}$$

$$f(x) = \sum_k^n [a_k x + b_k] U_g(x - k)$$

$$\int_0^x f(x') dx' = \int_0^x \sum_k^n [a_k x' + b_k] U_g(x' - k) dx'$$

$$= \sum_k^n \int_0^x [a_k x' + b_k] U_g(x' - k) dx'$$

$$= \sum_k^n \int_0^x [a_k x' U_g(x' - k) dx'] + \int_0^x [b_k U_g(x' - k) dx']$$

$$= \sum_k^n \int_0^x [a_k x' U_g(x' - k) dx'] + \int_0^x [b_k U_g(x' - k) dx']$$

$$\int_0^x U_g(x') dx' = \begin{cases} x, & se\ 0 \leq x \leq g \\ g, & caso\ contrário \end{cases}$$

$$\int_0^x x U_g(x') dx' = \begin{cases} \int_0^x x' dx', & \text{se } 0 \leq x \leq g \\ \int_0^g x' dx' + \int_g^x 0 dx', & \text{caso contrário} \end{cases}$$

$$= \begin{cases} \frac{x^2}{2}, & \text{se } 0 \leq x \leq g \\ \frac{g^2}{2}, & \text{caso contrário} \end{cases}$$

Obs: nesse ponto fazer $a_k x = a_k(x' + k - k)$ facilita muito a implementação computacional, logo,

$$\int_0^x f(x') dx' = \sum_k^n \int_0^x a_k(x'+k-k) U_g(x'-k) dx' + \sum_k^n \int_0^x b_k U_g(x'-k) dx'$$

$$= \sum_k^n \int_0^x a_k(x'-k) U_g(x'-k) + k U_g(x'-k) dx' + \sum_k^n \int_0^x b_k U_g(x'-k) dx'$$

$$= \sum_k^n \int_0^x a_k(x'-k) U_g(x'-k) dx' + k \int_0^x U_g(x'-k) dx' + \int_0^x b_k U_g(x'-k) dx'$$

$$\int_0^x \int_0^{x'} f(x'') dx'' dx'$$

$$= \sum_k^n \int_0^x \int_0^{x'} a_k(x''-k) U_g(x''-k) dx'' dx'$$

$$+ k \int_0^x \int_0^{x'} U_g(x''-k) dx'' dx' + \int_0^x \int_0^{x'} b_k U_g(x''-k) dx'' dx'$$

$$\int_0^x \int_0^{x'} U_g(x'') dx'' dx' = \begin{cases} \frac{x^2}{2}, & \text{se } 0 \leq x \leq g \\ \frac{g^2}{2} + g \int_g^x 1 dx', & \text{caso contrário} \end{cases}$$

$$= \begin{cases} \frac{x^2}{2}, & \text{se } 0 \leq x \leq g \\ \frac{g^2}{2} + g(x-g), & \text{caso contrário} \end{cases}$$

$$= \begin{cases} \frac{x^2}{2}, & \text{se } 0 \leq x \leq g \\ gx - \frac{g^2}{2}, & \text{caso contrário} \end{cases}$$

$$\int_0^x \int_0^{x'} x'' U_g(x'') dx' dx'' = \begin{cases} \int_0^x \frac{x'^2}{2} dx', & \text{se } 0 \leq x \leq g \\ \int_0^g \frac{x'^2}{2} dx' + \int_g^x \frac{g^2}{2} dx', & \text{caso contrário} \end{cases}$$

$$= \begin{cases} \frac{x'^3}{6}, & \text{se } 0 \leq x \leq g \\ \frac{g^3}{6} + \frac{g^2}{2}(x-g), & \text{caso contrário} \end{cases}$$

$$= \begin{cases} \frac{x'^3}{6}, & \text{se } 0 \leq x \leq g \\ -\frac{1}{3}g^3 + \frac{1}{2}g^3x, & \text{caso contrário} \end{cases}$$

Após o cômputo dessas integrais e determinação das constantes K_0, K_1 são mostradas as funções de base nas figuras da tabela 4-3 prontas para serem utilizadas em qualquer tipo de cálculo de estrutura eletrônica:

Tabela 4-3 funções de bases construídas a partir das wavelets da família Dabauchies

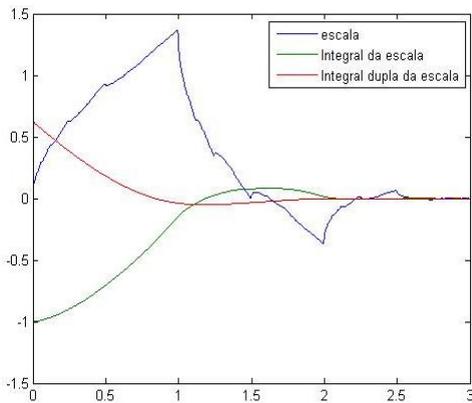


Figura 4-4 Wavelet DB2, Integral da DB2 e Integral dupla da DB2 (base proposta)

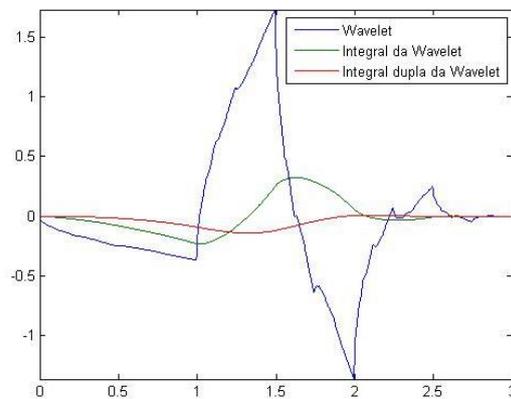


Figura 4-5 Wavelet DB2, Integral da DB2 e Integral dupla da DB2 (base proposta)

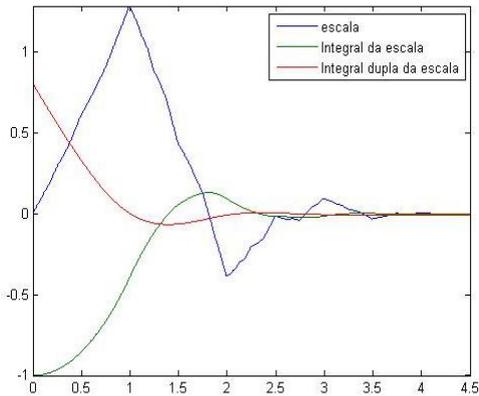


Figura 4-6 Wavelet DB2, Integral da DB2 e Integral dupla da DB2 (base proposta)

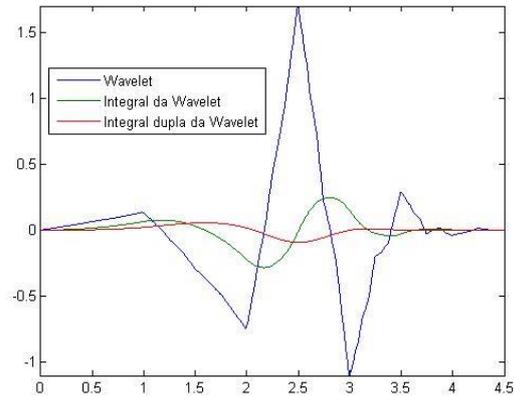


Figura 4-7 Wavelet DB2, Integral da DB2 e Integral dupla da DB2 (base proposta)

4.4.

Polinômios de Laguerre Generalizados, uma alternativa

Nesta seção é proposta a utilização da família de polinômios de Laguerre Generalizados. Os polinômios de Laguerre chamam atenção por dois motivos, o primeiro é que podem facilmente construir bases de Hilbert para espaço $L^2(0, \infty)$, isto é bases ortogonais, e o segundo se refere às facilidades de serem construídos, uma vez que podem ser definidos de forma recursiva. Não será aqui discutido profundamente os detalhes relacionados com os polinômios de Laguerre, mas sim o fundamental para o entendimento de seu uso. Diversos livros textos contém informações detalhadas sobre as famílias de polinômios de Laguerre, para mais detalhes em (STOER & BURLIRSCH,1980).

Os polinômios de Laguerre contribuem de forma significativa para a redução do esforço computacional, no cômputo da energia dos sistemas quânticos aqui trabalhados, isto pela facilidade de construção e pela sua ortogonalidade.

Os polinômios de Laguerre Generalizados de grau n podem ser obtidos através da equação 4-21

$$L_n^\alpha(x) = \sum_{i=0}^n (-1)^i \binom{n+\alpha}{n-i} \frac{x^i}{i!}. \quad 4-21$$

Os polinômios de Laguerre são ortogonais na semi-reta quando considera a função peso $x^\alpha e^{-x}$, conforme presente em (STOER & BURLIRSCH,1980). Temos então,

$$\int_0^{\infty} L_n^\alpha(x) L_m^\alpha(x) x^\alpha e^{-x} dx = \frac{\Gamma(n + \alpha + 1)}{n!} \delta_{n,m}. \quad 4-22$$

Conforme dito anteriormente os esses polinômios possuem a seguinte propriedade de recorrência

$$L_n^\alpha(x) = \left(2 + \frac{\alpha - 1 - x}{n}\right) L_{n-1}^\alpha(x) - \left(1 + \frac{\alpha - 1}{n}\right) L_{n-2}^\alpha(x). \quad 4-22$$

Além disso, também é útil para o desenvolvimento dos cálculos futuros as seguintes relações

$$\frac{d^k}{dx^k} L_n^\alpha(x) = (-1)^k L_{n-k}^{\alpha+k}(x) \quad 4-23$$

$$\frac{d^k}{dx^k} x^\alpha L_n^\alpha(x) = \binom{n + \alpha}{k} x^{\alpha-k} L_n^{\alpha-k}(x) \quad 4-24$$

Os polinômios de Laguerre não formam uma base ortonormal para o espaço $L^2(0, \infty)$ na métrica usual, uma vez que nem se quer pertencem a esse espaço, pois não possuem a integral do seu quadrado finita, mas eles podem ser úteis para a construção de uma base. A seguir, apresenta-se uma das bases para $L^2(0, \infty)$ equação 4-25.

$$e_n^\alpha(x) = \sqrt{\frac{n!}{\Gamma(n + \alpha + 1)}} x^{\frac{\alpha}{2}} L_n^{\alpha-k}(x) e^{-\frac{x}{2}} \quad 4-25$$

Fazendo $\alpha = 2$, podemos a partir de algumas manipulações algébricas construir uma nova base ortonormal para $L^2(0, \infty)$, diferente da usual apresentada pela equação 4-22, através da inserção de um parâmetro não linear. A inserção desse parâmetro não linear implica na possibilidade de variar a base utilizada como a base na série de Fourier das funções presentes em $L^2(0, \infty)$. A base em questão torna-se

$$q_n(x, a) = \sqrt{\frac{n!}{a^3 \Gamma(n + 3)}} x L_n^2\left(\frac{x}{a}\right) e^{-\frac{x}{a}}. \quad 4-26$$

Utilizando as base definida por $q_n(x)$ podemos aproximar qualquer função $\psi(x)$ em $L^2(0, \infty)$ pelo o uso das séries de Fourier conforme a equação 4-27

$$\sum_{n=1}^{\infty} q_n(x, a) \int_0^{\infty} q_n(x, a) \psi(x) dx \quad 4-27$$

Com uma modificação sutil, conforme visto na equação 4-26, é possível introduzir uma nova base para $L^2(0, \infty)$, dessa vez não de Hilbert, mas de Hamel, isto é um conjunto de funções que seja linearmente independente, mas que esteja apto a ser utilizado para aproximar qualquer função em $L^2(0, \infty)$.

$$q'_n(x, a) = \sqrt{\frac{n!}{a^3 \Gamma(n+3)}} L_n\left(\frac{x}{a}\right) e^{-\frac{x}{a}} \quad 4-28$$

Proposição 4.4.

$$\bigvee_i q'_n(x, a) = L^2(0, \infty) \quad \forall a \in \mathcal{R}^+.$$

Dessa forma, foi criado um método para realizar a variação da base utilizada na serie de Fourier, possibilitando achar uma base mais adequada para aproximação de um ponto qualquer definido no $L^2(0, \infty)$, possibilitando explorar um subespaço, “mais correto”, no sentido em que podem ser necessários muitos elementos de base para aproximar um dado ponto em $L^2(0, \infty)$, utilizando a base errada, muito embora esta ainda seja base para o $L^2(0, \infty)$.

Apesar dos polinômios de Laguerre modificados formarem uma base, $q'_n(x, a)$ para o $L^2(0, \infty)$, imediatamente surge a desvantagem relativa a quantidade de funções de x presentes em cada polinômio, essa quantidade é dada por $n + 1$, onde n é grau do polinômio utilizado, o que pode representar um alto custo computacional.

Em geral, para o cômputo dos termos de energia cinética e potencial de um sistema de dois elétrons, estão presentes equações no seguinte formato

$$\int f K f, \quad 4-29$$

ou

$$\int f K g, \quad 4-30$$

onde, tanto f como g são funções desconhecidas em $L^2(0, \infty)$, as quais se deseja aproximar, e K é um operador $K: L^2(0, \infty) \rightarrow L^2(0, \infty)$.

Aqui então, do ponto de vista computacional, são vislumbradas duas hipóteses para realizar o cálculo dos funcionais definidos pelas equações 4-29 e 4-30. A seguir as possibilidades

- a) Realiza-se o cálculo da integral da série de Fourier de f ou g ;
- b) Realiza-se a soma parciais das integrais dos termos da série de Fourier de f ou g ;

Para ilustrar a hipótese a) temos:

$$\int \left(\sum_{i=1}^n c_i \varphi_i \right) K \left(\sum_{i=1}^n c_i \varphi_i \right) d\mu \quad 4-29$$

Para ilustrar a hipótese b) temos:

$$\sum_{i=1}^n \sum_{j=1}^n c_i c_j \int_0^{\infty} \varphi_i K \varphi_j d\mu. \quad 4-30$$

As duas abordagens resultam exatamente no mesmo número. No entanto, na alternativa a) é realizada apenas uma única integral, em contrapartida, na abordagem b) a seguinte quantidade de integrais dada por n^2 . Então, nessa tese é sugerida e empregada nos cálculos sempre a alternativa b).

A grande vantagem em se utilizar elementos de base definidos por $q'_n(x, a)$ é seu caráter recursivo, isto é, pode definir o termo $q'_{n+1}(x, a)$ através do uso das relações de recorrência já previamente expostas. Além disso, calcular as quantidades mencionadas, é uma questão de aplicar alguma regra de quadratura, uma vez que é possível avaliar as funções em todos os pontos da regra de quadratura de forma muito eficiente, graças ao caráter recursivo da base proposta

5

FUNDAMENTOS DE OTIMIZAÇÃO EVOLUCIONÁRIA

Muitas das teorias clássicas de otimização foram primeiramente motivadas por problemas físicos, e têm relação com grandes matemáticos: Gauss, Lagrange, Euler, Bernoulli, dentre outros. Durante os desenvolvimentos recentes de problemas de otimização para a tomada de decisão, as técnicas clássicas têm sido reexaminadas, estendidas, e algumas vezes redescobertas. Novos *insights* foram desenvolvidos e novas técnicas têm sido desenvolvidas, como por exemplo, Algoritmos Evolucionários.

5.1.1.

Fundamentos de programação não-linear

Definição 5.1 (Ponto de mínimo)

Seja uma função $f: S \rightarrow \mathcal{R}$, seja $x^*, y \in S$, x^* será um ponto de mínimo de f se existir uma região com raio ε em torno de x^* tal que

$$f(x^* + y) > f(x^*) \quad \forall |y| < \varepsilon \quad 5-1$$

A função f é chamada de função objetivo. O mínimo de uma função é chamado local quando consiste no menor valor da função em uma pequena vizinhança. Quando se trata do menor valor em toda uma região de interesse, esse mínimo é chamado de mínimo global. Esse é, na maioria dos casos, o objeto de estudo. Quando S contém um grande número de dimensões e f é não-linear, pode ser extremamente difícil encontrar o mínimo global, devido à presença de uma grande quantidade de mínimos locais na região em questão. A grande dificuldade associada está na especificação de um critério que identifique mínimos com a propriedade de ser global, pois todo o mínimo global também é um mínimo local, e uma vez que não são previamente conhecidos todos os mínimos locais de uma dada região, em muitos dos casos, pode ser falsa a identificação de globalidade em um mínimo local especificado. Uma função que contém mínimos locais e globais é mostrada na Figura 5-1.

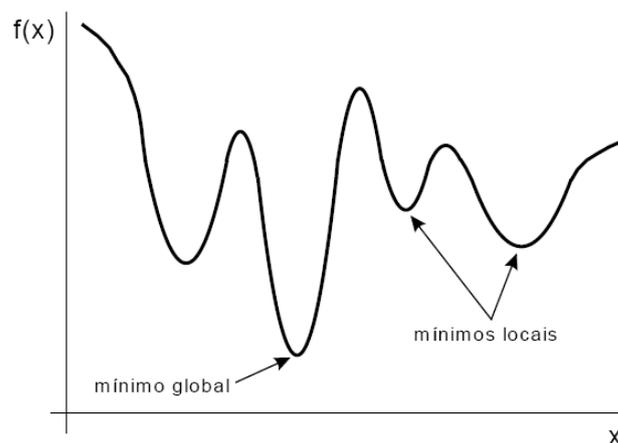


Figura 5-1 Ilustra uma função que possui mínimos locais e globais.

5.1.2.

Modelos determinísticos e estocásticos

Como dito na seção anterior, pode não ser simples identificar mínimos globais. Dessa forma, quando a questão principal é a localização de mínimos globais, é recomendável que sejam utilizados métodos pouco sensíveis às soluções iniciais. Esses métodos são chamados de métodos de otimização global e podem ser determinísticos ou estocásticos. Na Figura 5-2 é mostrada uma taxonomia dos métodos de otimização globais determinísticos e estocásticos.

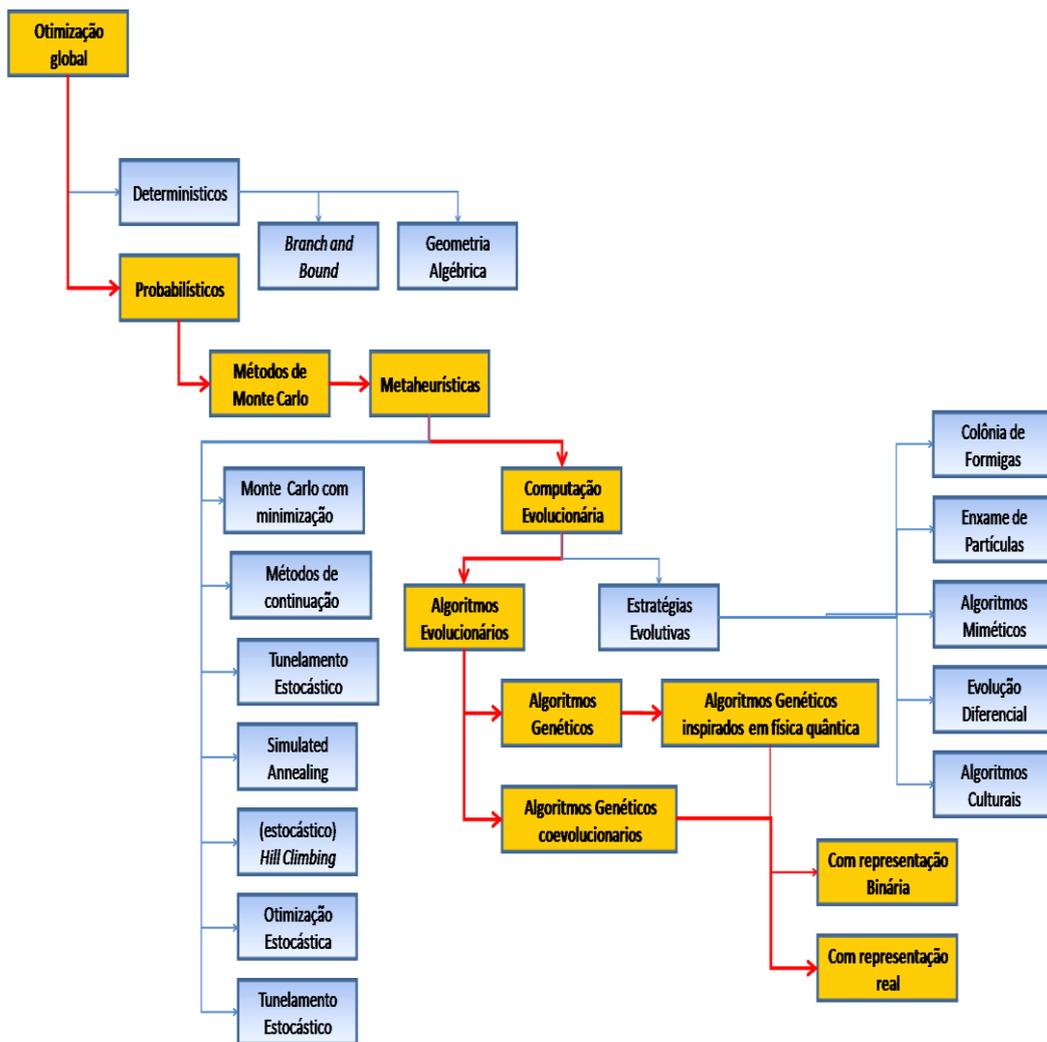


Figura 5-2 Taxonomia dos métodos de otimização global, determinísticos e estocásticos.

Neste trabalho, são particularmente importantes os Algoritmos Evolucionários. Tais algoritmos têm sido extensivamente utilizados na literatura para resolver problemas classificados como difíceis. Os algoritmos evolutivos ganharam destaque quando apresentados pela primeira vez, em Michigan por John Holland em 1990.

John Holland criou um algoritmo de otimização inspirado em um processo biológico com argumentos matemáticos consistentes, a teoria dos *schemata* e dos blocos construtores (MICHALEWICZ, 1994). Tais algoritmos, chamados de Algoritmos Genéticos (AGs), são inspirados no princípio Darwiniano da evolução das espécies e na genética. Esses algoritmos são probabilísticos e fornecem um mecanismo de busca paralela e adaptativa baseado no princípio de sobrevivência dos mais aptos e na reprodução (MICHALEWICZ, 1994). Eles têm sido usados

com sucesso para encontrar soluções adequadas para problemas de inúmeras áreas da ciência e da engenharia (MICHALEWICZ, 1994) e deram origem a um novo ramo científico: a Computação Evolucionária.

Apesar de a computação evolucionária ser comprovadamente eficaz na busca por mínimos globais (MICHALEWICZ, 1994), em sua forma básica ainda apresentam problemas que a impossibilita atuar em certas funções. Por exemplo, a quantidade de avaliações de alternativas dentro do conjunto de soluções pode ter uma alta demanda computacional e de memória. Além disso, AGs também podem ficar presos em mínimos locais do espaço de busca, portanto podem não encontrar a solução com melhor qualidade para o problema abordado.

Essa classe de algoritmos sofreu várias adaptações e algumas das barreiras referentes à sua utilização foram transpostas. Nas próximas seções são mostrados os aspectos práticos referentes à utilização de mecanismos evolucionários.

5.1.3.

Fundamentos de programação evolucionária

Algoritmos genéticos

Algoritmos Genéticos (AGs) pertencem à classe dos Algoritmos Evolutivos (AEs) (FUKUDA,1996), técnicas de otimização inspiradas pela observação de fenômenos naturais. Algoritmos que tentam imitar o princípio da seleção natural são denominados AGs.

Esses algoritmos possuem algumas características que favorecem as suas aplicações, o paralelismo implícito e, geralmente, a sensibilidade da busca pelo ótimo global não é afetada pela arbitrariedade de condições iniciais.

Resumidamente, é possível explicar o funcionamento de um algoritmo genético clássico definindo-se alguns conceitos básicos. O primeiro passo é gerar uma população inicial, cujos indivíduos representam possíveis soluções para um determinado problema. Essa população inicial pode ser gerada a partir de valores aleatórios ou a partir de valores predefinidos (sementes). Cada indivíduo é avaliado de acordo com o problema em questão, e os mais aptos são mantidos enquanto os demais são eliminados. Por meio de operadores genéticos (cruzamento e mutação), os indivíduos restantes geram descendentes (reprodução), os quais tem uma grande possibilidade de serem mais aptos do que

seus genitores. A reprodução é repetida até que um critério de parada seja satisfeito. Essa condição pode estar relacionada a uma solução satisfatória, que pode ser o número de gerações ou até mesmo o tempo de processamento.

5.1.4.

Operadores e Parâmetros Genéticos

Os operadores genéticos têm a função de, por meio de um processo iterativo, transformar a população inicial em uma população que contenha um resultado satisfatório. Um algoritmo genético clássico é composto de três operações:

1. Seleção;
2. Cruzamento;
3. Mutação.

A idéia básica da reprodução é selecionar os melhores indivíduos da população corrente através de uma função de aptidão. Os indivíduos com um alto valor de aptidão terão uma maior probabilidade de contribuir com um ou mais descendentes na próxima geração.

A operação de reprodução pode ser executada de várias maneiras, porém, o método mais utilizado é o método da roleta. Nesse método, cada indivíduo da população tem probabilidade de ser selecionado proporcional ao seu valor de aptidão, calculado pela função custo. A Figura 5-3 ilustra o esquema de evolução dos AGs Clássicos.

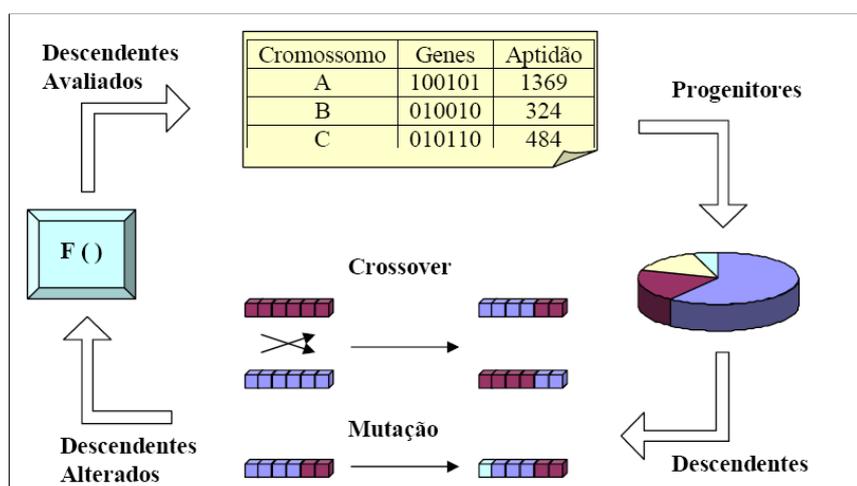


Figura 5-3 Esquema de evolução dos AGs Clássicos.

5.1.5.

Mecanismo evolutivo em detalhes

A criação de uma população P_0 consiste de uma amostragem aleatória de elementos, conforme algum procedimento previamente definido, de um conjunto D . Denota-se $M(D)$ um espaço de probabilidade de um domínio D . A distribuição de probabilidade $h_0 \in M(D)$ é usualmente uniforme.

$$h_0(A) = \frac{\mu(A)}{\mu(D)}, \text{ onde } A \subset D \quad 5-2$$

A princípio nenhuma região A do domínio D é preferida, daí $h_0(A)$ ser uma distribuição uniforme. Isso porque não há informação alguma sobre pontos de mínimos globais no início da busca. Quando a busca se inicia, h_0 pode se concentrar em algumas regiões de D .

A avaliação da população de amostras aleatórias, P_t , é feita através do computo dos valores da função objetivo para todos os elementos de P_t , i.e. $\{\phi(x)\}_t, t = 0, 1, 2, \dots \forall x \in P_t$.

A modificação das amostras é feita substituindo cada elemento x de P_t de acordo com algum critério específico. Em geral, as modificações das amostras em P_t consistem em reter e modificar alguns pontos de amostras prévias $\{P_i\}, i = 0, 1, \dots, t - 1$.

Para construir a nova população P_{t+1} são utilizados os operadores genéticos, crossover e mutação, em populações anteriores P_t ou amostrados de P_t de acordo com uma distribuição de probabilidade já conhecida.

Seja $h_{t+1} \in M(D)$ a distribuição de probabilidade de se sortear um ponto de D na população P_{t+1} . As estratégias para a atualização de h_{t+1} é que vão determinar comportamento e eficiência de algoritmo evolutivo.

Definição 5.3

Um ponto $x^* \in \mathcal{R}^n$ é dito mínimo global de uma função objetivo, $\phi: \mathcal{R}^n \rightarrow \mathcal{R} \in \Gamma$, onde Γ é um espaço linear de funções se

$$\phi(x^*) \leq \phi(x), \forall x \in \mathcal{R}^n \quad 5-3$$

Definição 5.4

Uma estratégia estocástica de otimização global é dita assintoticamente correta, se e somente se está apta a encontrar maximizadores/minimizadores globais com probabilidade 1 depois de um número infinito de interações.

Apesar da definição acima, para fins de computação nesse trabalho admite-se um número infinito de interações como um número arbitrariamente grande.

Seja x^* um mínimo global de uma função $\phi: \mathcal{R}^n \rightarrow \mathcal{R}$ e considere a seguinte sequência $\{Y_t = \phi(TP_t)\}_{t=0,1,2,\dots}$, onde T é o mapeamento definido pelo conjunto de operações determinadas pela heurística evolutiva em P_t .

Definição 5.5

Diz-se que uma estratégia de otimização global é globalmente convergente se a sequência $\{Y_t\}, t = 1, 2, \dots$ onde $\{Y_t = \phi(TP_t)\}_{t=0,1,2,\dots}$, onde T é o mapeamento definido pelo conjunto de operações determinadas pela heurística evolutiva em P_t , converge para $\phi(x^*)$, no seguinte sentido

$$\forall \varepsilon > 0, \left(\lim_{t \rightarrow \infty} \sum_{i=0}^t \text{Prob}(|\phi(x^*) - Y_i| > \varepsilon) \right) < \infty. \quad 5-4$$

Teorema 5.6

Se um algoritmo evolutivo é uma estratégia globalmente convergente, então

$$\forall \varepsilon > 0, \left(\lim_{t \rightarrow \infty} \text{Prob}(|\phi(x^*) - Y_t| < \varepsilon) \right) = 1. \quad 5-5$$

Prova.

Seja $a_i \in \mathcal{R}_+, \forall i \in \mathbb{N}$, fazendo $a_i := \text{Prob}\{|\phi(x^*) - Y_i| > \varepsilon\}, \forall \varepsilon > 0$, para todo evolutivo que seja globalmente convergente $\sum_{i=0}^{\infty} a_i < \infty$, pela definição 5.5. De acordo com o resultado

$$\sum_{j=0}^{\infty} a_j < \infty \Rightarrow a_j \rightarrow 0, \quad 5-6$$

Prova Ver (Kubrusly, 2001).

tem-se,

$$\forall \varepsilon > 0, \sum_{t=0}^{\infty} \text{Prob}\{\phi(x^*) - Y_t \geq \varepsilon\} < \infty \Rightarrow \text{Prob}\{|\phi(x^*) - Y_t| \geq \varepsilon\} \rightarrow 0, \quad 5-7$$

que é equivalente a

$$\forall \varepsilon > 0, \left(\lim_{t \rightarrow \infty} \text{Prob}(|\phi(x^*) - Y_t| < \varepsilon) \right) = 1. \quad 5-8$$

5.1.6.

Algoritmos Evolutivos com Inspiração Quântica

Com o ideal de melhorar o desempenho dos AGs, surgiram vários trabalhos buscando inspiração em muitos paradigmas como, por exemplo, Computação Quântica, para reduzir o esforço computacional dos algoritmos clássicos de otimização.

Pesquisas para a fusão de Algoritmos Evolucionários com a Computação Quântica têm sido feitas desde o final da década de 90 (NARAYANAN & MOORE, 1996), (HEY, 1996) e deram origem aos Algoritmos Genéticos com Inspiração Quântica que são novos e promissores algoritmos. Trabalhos inspirados em computação quântica são desenvolvidos em (CRUZ, 2005), (CRUZ, 2007), (NARAYANAN & MOORE, 1996) e (HAN & KIM, 2000). Entretanto, estes não contemplam soluções com representação direta no espaço dos números reais.

Em (CRUZ, 2007) foi desenvolvido um Algoritmo Evolucionário com Inspiração Quântica para Resolução de Problemas de Representação Real (*AIEQ*) que se mostrou superior à grande maioria dos Algoritmos Evolucionários existentes, nos casos estudados.

5.2.

Algoritmo Evolucionário com Inspiração Quântica para Resolução de Problemas de Representação Real

A primeira abordagem sobre uma representação específica para problemas numéricos em conjuntos reais foi feita em (CRUZ, 2007). Essa representação simula a superposição de estados da Mecânica Quântica através de funções de densidade de probabilidade, representando todos os possíveis estados que uma variável contínua possa assumir em um domínio. Essa abordagem, que possui paralelismo quântico, mostrou-se de notável eficácia para resolução de problemas de otimização no domínio dos reais. A Figura 5-4 mostra o pseudocódigo do Algoritmo Genético com Inspiração Quântica.

```

iniciar
1.  $t \leftarrow 1$ 
2. Gerar população quântica  $Q(t)$  com  $N$  indivíduos com  $G$  genes
3. enquanto ( $t \leq T$ )
4.    $E(t) \leftarrow$  gerar indivíduos clássicos usando indivíduos quânticos
5.   se ( $t=1$ ) então
6.      $C(t) \leftarrow E(t)$ 
7.   senão
8.      $E(t) \leftarrow$  recombinação entre  $E(t)$  e  $C(t)$ 
9.     avaliar  $E(t)$ 
10.     $C(t) \leftarrow K$  melhores indivíduos de  $[E(t) + C(t)]$ 
11.  fim se
12.   $Q(t+1) \leftarrow$  Atualiza  $Q(t)$  usando os  $N$  melhores indivíduos de  $C(t)$ 
13.   $t \leftarrow t + 1$ 
14. fim enquanto
fim

```

Figura 5-4 Pseudocódigo do AEIQ-R.

Além de oferecer uma representação mais adequada para problemas cujo domínio é o conjunto dos *Reais*, o modelo fornece as seguintes vantagens em relação aos AGs tradicionais e aos que possuem inspiração quântica mas que operam sobre espaços binários, os *AEIQ-B* (NARAYANAN & MOORE, 1996):

- *Menor tempo para convergência;*
- *O conhecimento sobre o problema que está sendo otimizado é armazenado diretamente nos cromossomos quânticos;*
- *Capacidade de convergir de forma eficiente, mesmo com populações com poucos indivíduos.*

A população quântica

A população quântica do algoritmo é formada por um conjunto de funções de densidade de probabilidade, ou seja, um vetor de variáveis aleatórias, especificadas segundo uma distribuição qualquer. A equação (5-9) representa um indivíduo de uma população quântica e a Figura 5-5 mostra graficamente a representação de genes quânticos de um indivíduo.

$$Q_i = [p_{ij}, p_{ij+1}, p_{ij+2}, \dots] \quad 5-9$$

onde $i = 1, 2, 3, \dots, N$, $j = 1, 2, 3, \dots, N$, e as funções p_{ij} representam as funções densidade de probabilidade. Estas funções densidade de probabilidade são usadas pelo *AEIQ* para gerar os valores para os genes dos indivíduos clássicos.

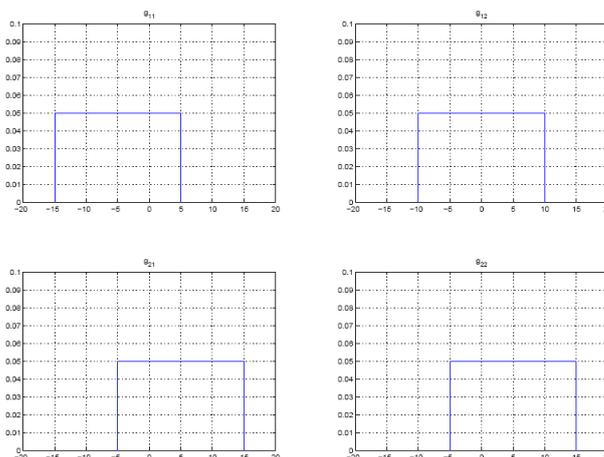


Figura 5-5 Representação de um indivíduo composto por funções de densidade de probabilidade uniformes.

Observação dos indivíduos clássicos a partir de indivíduos quânticos

A partir do conjunto de genes quânticos são observados um ou mais indivíduos através do sorteio de um número aleatório. Esse número é obtido partir de uma distribuição de probabilidade que compõe o indivíduo quântico. O sorteio é feito conforme a sequência de procedimentos a seguir: é sorteado um número aleatório entre zero e um, posteriormente é calculada a imagem inversa da função de densidade acumulada no ponto sorteado, esse corresponde a um gene clássico de um indivíduo que compõe a população clássica. A Figura 5-6 mostra a função de densidade acumulada para o gene representado pela densidade uniforme e como ele organizar-se-ia no indivíduo quântico apresentado na Figura 5-5 para o caso implementado em (CRUZ, 2007).

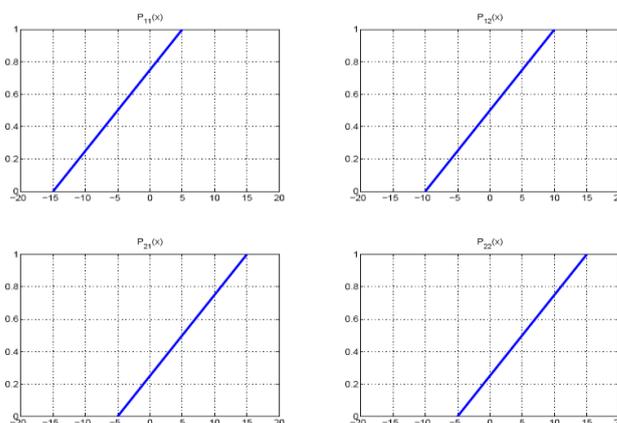


Figura 5-6 Funções de probabilidade acumulada associadas a cada gene de um indivíduo quântico

A variável aleatória uniforme, algumas vezes referida no texto como pulso quadrado, descrita pela equação (6-10) é uma das funções mais simples de ser implementadas computacionalmente.

$$p_{ij}(x) = \begin{cases} \frac{1}{U_{ij}-L_{ij}} & \text{se } L_{ij} \leq x \leq U_{ij} \\ 0 & \text{caso contrário} \end{cases} \quad 5-10$$

onde L_{ij} é o limite inferior e U_{ij} o limite superior do intervalo no qual o gene j do i -ésimo indivíduo quântico pode assumir valores quando observado.

A função de densidade acumulada de uma variável aleatória uniforme é dada pela seguinte equação,

$$x_{mj} = r_{mj} * \sigma_{i2} + \left(\mu_{ij} - \frac{\sigma_{ij}}{2} \right) \quad 5-11$$

Para o caso em que P_{ij} é um pulso quadrado, pode-se representar o gene quântico armazenando-se os valores dos limites inferior e superior para cada gene, ou armazenando a posição do ponto central do pulso quadrado e a largura do mesmo. Por exemplo, considerando um indivíduo quântico Q_{ij} formado por dois pulsos quadrados, e supondo que estes dois pulsos têm uma largura igual a dois e estão posicionados de tal modo que o seu centro está localizado nas posições -0.5 e 0.5 , respectivamente, o cromossomo quântico pode ser representado, caso se esteja usando largura e centro com esses valores para os genes, por $Q_i = [\mu_{i1} = -0.5, \mu_{i2} = 0.5, \sigma_{i1} = 2, \sigma_{i2} = 2]$, onde μ_{i1} e μ_{i2} indicam o centro e σ_{i1} e σ_{i2} representam a largura dos dois pulsos quadrados respectivamente. A altura desses pulsos deve ser calculada de modo que a propriedade de normalização da função de densidade de probabilidade formada pela soma desses pulsos quadrados seja respeitada. Assim, a altura dos pulsos quadrados deve ser tal que a área total dos mesmos seja igual a um, e pode ser calculada usando-se a equação 6-1. No exemplo dado, cada pulso terá, portanto, uma altura igual a 0.5 . Ao usar o pulso quadrado, é possível utilizar, pelo menos, duas estratégias diferentes de inicialização dos mesmos: na primeira, criam-se pulsos quadrados com uma largura igual a $U_{ij} - L_{ij}/N$ (onde N é o número de indivíduos quânticos) e com seus centros distribuídos uniformemente ao longo de todo o domínio das variáveis; na segunda, criam-se todos os pulsos com o limite inferior

e superior igual ao limite inferior e superior do domínio, respectivamente, sob o qual se deseja otimizar a função objetivo. Foi escolhida a segunda estratégia e exemplificado o seu uso a partir da seguinte situação. Considere um problema com duas variáveis x e y , ambas no intervalo $[-100, 100]$. A função de densidade de probabilidade é representada por um pulso quadrado e a população quântica inicial $Q(0)$ tem cinco indivíduos quânticos. É possível representar os indivíduos quânticos que formam a população inicial conforme a Equação (6-12)

$$Q(0) = \begin{bmatrix} q_1 = [(\mu = 0, \sigma = 200; (\mu = 0, \sigma = 200))] \\ q_2 = [(\mu = 0, \sigma = 200; (\mu = 0, \sigma = 200))] \\ q_3 = [(\mu = 0, \sigma = 200; (\mu = 0, \sigma = 200))] \\ q_4 = [(\mu = 0, \sigma = 200; (\mu = 0, \sigma = 200))] \\ q_5 = [(\mu = 0, \sigma = 200; (\mu = 0, \sigma = 200))] \end{bmatrix} \quad 5-12$$

Resumidamente, a observação de indivíduos clássicos a partir de quânticos é feita da seguinte forma:

1. Gera-se um número aleatório r no intervalo $[0,1]$;
2. Identifica-se o ponto x , dado que $P_{ij}(x) = \int p_{ij} dx$, $x = P_{ij}^{-1}(r)$ onde, P_{ij} é a função de densidade de probabilidade acumulada, P_{ij}^{-1} é a sua inversa;
3. x é o valor observado para o gene j do indivíduo clássico i .

A atualização da população clássica e quântica

Todos os indivíduos clássicos são construídos seguindo os passos descritos na seção anterior dando origem a um conjunto chamado aqui de população clássica. Os indivíduos da população nova são avaliados (passo nove do algoritmo da Figura 5-4) pela função de avaliação que se deseja otimizar. Apesar de ser possível, não é, provavelmente, conveniente usar operações de mutação nos indivíduos novos ou antigos, já que os indivíduos quânticos, por si só, já introduzem um efeito aleatório nas populações sendo evoluídas. Com a população clássica gerada, deve-se tomar uma decisão sobre como substituir os indivíduos de uma eventual população pré-existente (passo 10 do algoritmo da Figura 5-4). De fato, a estratégia de substituição dos indivíduos da população da geração anterior pelos indivíduos criados na nova geração deve ser usada a todo o momento, com exceção da primeira geração, onde não existe uma população anterior. Existem várias abordagens possíveis para essa etapa do algoritmo:

1. Substituir todos os indivíduos da população antiga pelos indivíduos da população nova (equivalente a ausência de elitismo e *steady-state* dos AGs convencionais);
2. Substituir todos os indivíduos da população antiga pelos indivíduos da população nova, mas preservando o melhor indivíduo da nova população (equivalente ao elitismo dos AGs tradicionais);
3. Substituir os n piores indivíduos da população antiga pelos n melhores indivíduos da população nova (equivalente ao *steady-state* dos AGs tradicionais);
4. Substituir os n piores indivíduos da população antiga por n indivíduos quaisquer da população nova.

Cada uma dessas abordagens apresenta vantagens e desvantagens. A primeira opção diminui o tempo de processamento, pois elimina a necessidade de ordenação dos indivíduos da população antiga e da população nova. No entanto, essa abordagem tende a introduzir muito ruído, já que os indivíduos da população antiga são completamente substituídos e a avaliação do melhor indivíduo pode piorar significativamente, dependendo do problema que se deseja otimizar. A opção dois resolve esse problema preservando o indivíduo mais apto da população clássica na população nova. A opção três, apesar de exigir a ordenação dos elementos da população antiga e da população nova, apresenta a alternativa mais conservadora de busca, sem provocar grandes alterações na população que está sendo evoluída. Finalmente, a opção quatro combina o conservadorismo da opção três, mantendo os indivíduos mais bem avaliados na população em evolução, mas diminuindo o tempo de processamento por não necessitar avaliar todos os indivíduos da população nova.

Para a atualização dos indivíduos da população quântica os procedimentos são dependentes do tipo de função de densidade de probabilidade que se definiu para representar os genes quânticos (CRUZ, 2007).

A atualização da população quântica consiste em alterar os momentos estatísticos das funções de densidade de probabilidades utilizadas para representar os genes quânticos usando uma heurística que tem os seguintes objetivos:

1. Reduzir o espaço de busca da função que se quer otimizar. No AEIQ, isto é feito reduzindo-se o tamanho da região onde a função

de densidade de probabilidade (genes quânticos) tem probabilidade diferente de zero;

2. Mapear as regiões mais promissoras do espaço de busca. Isto deve ser feito aumentando-se a probabilidade de se observar um determinado conjunto de valores para o gene clássico nas proximidades dos indivíduos mais bem sucedidos da população clássica.

Para o primeiro objetivo, pode-se usar um decaimento exponencial da largura dos pulsos quadrados, um decaimento linear ou, ainda, um processo similar ao usado em algoritmos de estratégia evolutiva, chamado “regra do 1/5” (MICHALEWICZ, 1994). A regra do 1/5 faz com que a modificação de largura dos pulsos seja executada de forma homogênea para todos os genes quânticos e para todos os indivíduos quânticos. A heurística usada para determinar se a largura do gene será aumentada ou diminuída é a seguinte (MICHALEWICZ, 1994): se menos de 20% da população clássica criada na geração atual tiver uma avaliação melhor do que na geração anterior, a largura do gene é reduzida; se esta taxa for maior do que 20%, a largura do gene é aumentada; caso a taxa seja exatamente igual a 20%, nenhuma alteração é feita. Matematicamente, isto pode ser representado pela equação (6-13).

$$\sigma_{ij} = \left\{ \begin{array}{ll} \sigma_{ij} * \delta & \text{se } \varphi < \frac{1}{5} \\ \frac{\sigma_{ij}}{\delta} & \text{se } \varphi > \frac{1}{5} \\ \sigma_{ij} & \text{caso contrário } \varphi = \frac{1}{5} \end{array} \right\}$$

5-13

onde σ_{ij} é a largura do j -ésimo gene do i -ésimo indivíduo quântico em $Q(t)$, δ é um valor arbitrário, em geral no intervalo $]0, 1[$ e φ é a taxa que indica quantos indivíduos da nova população clássica foram melhores do que os indivíduos da população clássica na geração anterior.

O argumento para o uso desta regra se baseia na seguinte heurística: se menos de 20% dos indivíduos da população clássica gerada tiverem melhorado de uma geração para outra, é provável que o algoritmo esteja fazendo a busca em uma região muito grande e, neste caso, pode ser interessante reduzir o espaço de busca; se mais de 20% dos indivíduos tiverem melhorado de uma geração para outra, é provável que o algoritmo esteja fazendo a busca em uma região muito

pequena (e, portanto, achando facilmente, indivíduos com avaliações melhores) e o mesmo deve tentar ampliar a região de busca; a taxa de 20% é considerada ideal e, portanto, nenhuma alteração é feita caso essa taxa seja medida.

Já no passo dois da atualização dos genes quânticos, o processo pode ser definido em termos de quais indivíduos da população clássica serão usados para atualizar os pulsos da população quântica. Podem ser escolhidos os indivíduos mais aptos, os indivíduos menos aptos, um conjunto aleatório de indivíduos ou usar um processo de roleta para selecionar quais os indivíduos que farão parte do conjunto de indivíduos clássicos que serão usados para atualizar os indivíduos quânticos. Após escolherem-se quais indivíduos serão usados (o número de indivíduos clássicos escolhidos deve ser igual ao número total de indivíduos quânticos e, portanto, o número de indivíduos clássicos não pode ser menor do que o número de indivíduos quânticos), o centro dos pulsos de cada indivíduo quântico deve ser modificado em relação ao valor de cada gene clássico. Em geral, pode-se usar um cálculo que desloque o centro dos pulsos na direção do ponto indicado pelo gene do indivíduo clássico. Por exemplo, supondo-se que o centro do gene quântico na geração t seja dado por $\mu_{ij}(t)$ e o valor do gene clássico seja dado por x_{ij} , então, pode-se calcular a nova posição do gene quântico na geração $t + 1$, através da equação $\mu_{ij}(t + 1) = \mu_{ij}(t) + \lambda(x_{ij} - \mu_{ij}(t))$, onde λ indica o percentual que se quer deslocar o centro do gene quântico na direção do gene clássico.

Execução do Algoritmo Evolutivo com Inspiração Quântica

Todos os passos do algoritmo descrito anteriormente devem ser repetidos por um número de gerações T , de acordo com o que foi mostrado na Figura 5-4. Além do pseudocódigo definido nessa figura, é possível definir um diagrama que mostra como as partes que constituem o modelo se relacionam entre si. Esse diagrama pode ser visto na Figura 5-7. Essa figura ilustra graficamente em ordem cronológica, os acontecimentos que ocorrem durante a evolução do algoritmo.

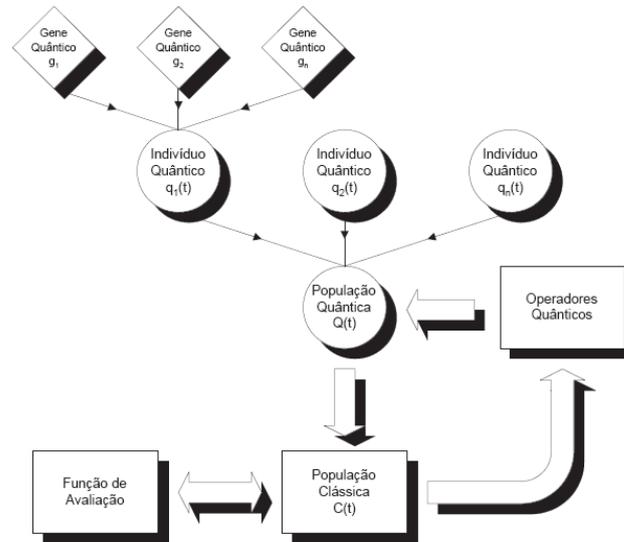


Figura 5-7 Diagrama completo do Sistema Evolutivo com Inspiração Quântica

O diagrama da Figura 5-7 mostra que a população quântica $Q(t)$ é composta de indivíduos quânticos que, por sua vez, são formados por genes quânticos. Essa população quântica é usada para gerar uma população clássica que é então avaliada e usada para modificar os indivíduos da população quântica em um processo iterativo.

6

FUNÇÕES DE ONDA EVOLUCIONÁRIAS E A SOLUÇÃO DA EQUAÇÃO DE SCHRÖDINGER

Neste capítulo é apresentado um conjunto de resultados, que fazem parte do arcabouço e ferramental necessários para o desenvolvimento desta tese. Algumas considerações são feitas sobre as soluções propostas e são descritos em detalhes quais os procedimentos gerais para se empregarem técnicas baseadas em Algoritmos Evolutivos para construir soluções válidas para a equação de Schrödinger. Em cada seção é apresentada uma breve descrição teórica sobre os métodos e artifícios utilizados e a forma de empregar a computação evolutiva em sinergia com tais métodos.

6.1.

A solução da equação Schrödinger por Algoritmos Evolucionários

A computação evolutiva ou evolucionária é a técnica que emprega algoritmos evolutivos para resolver problemas de otimização. Quando essa técnica é aplicada com a finalidade de realizar a construção de algo, tem-se a origem do projeto evolucionário ou evolutivo. Por isso, denomina-se, nesta tese, de Funções de Ondas Evolucionárias as funções de onda encontradas como soluções da equação de Schrödinger por métodos que empregam algoritmos evolucionários.

A equação de Schrödinger é uma equação de autovalor. Para muitas aplicações é procurada a autofunção do operador de Schrödinger associada ao menor autovalor desse operador. No entanto, como o operador de Schrödinger para sistemas moleculares e atômicos é um operador auto adjunto (TESCH,2009). Esse problema é equivalente a resolver um problema de otimização. Precisa-se encontrar o mínimo de um funcional conhecido como Range Numérico dado pela equação 6-1. Isso é justificado pelo teorema 6-1. Na figura 6-1 é mostrado o fluxograma da execução de um algoritmo evolutivo.

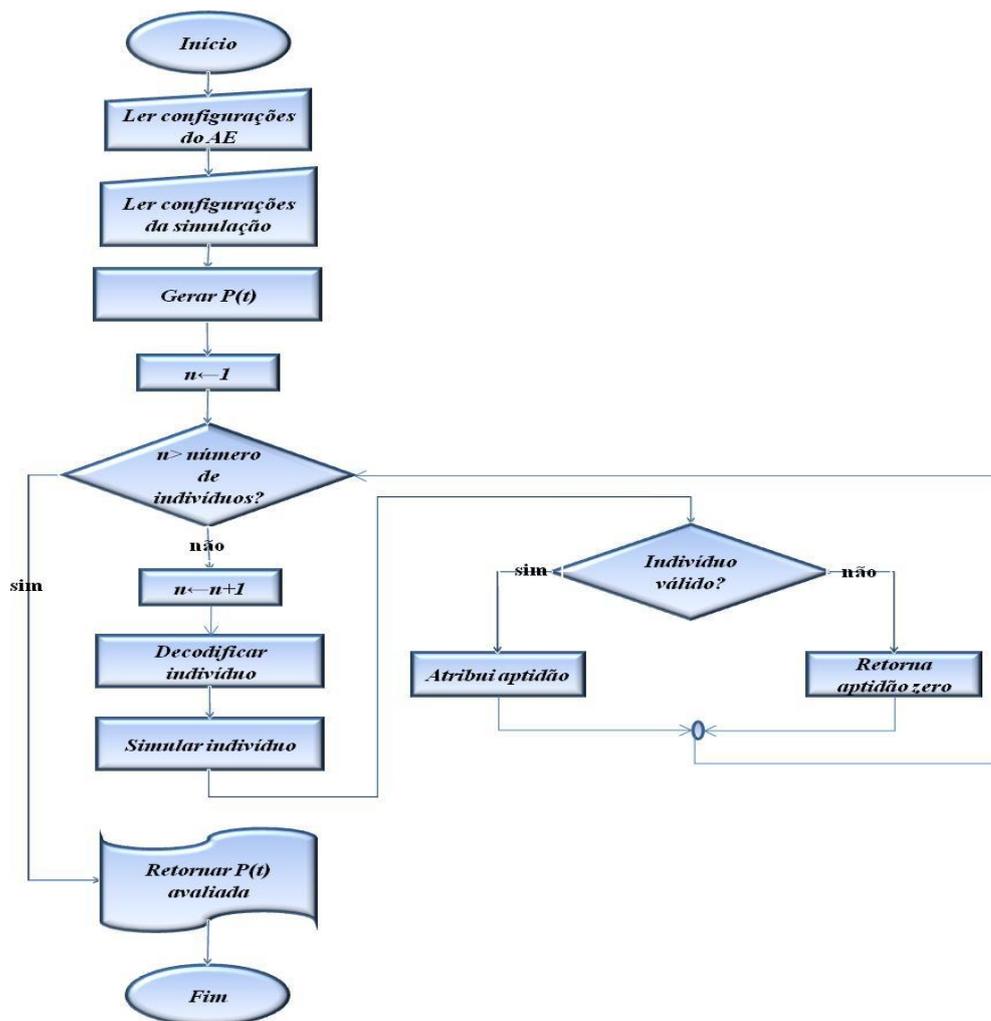


Figura 6-1 Fluxograma da execução de Algoritmos Evolucionários empregados na busca por funções de onda evolucionárias.

Teorema 6.1

Se Ψ é uma função que satisfaz as propriedades já discutidas então tem-se

$$\frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq \epsilon_0 \tag{6-1}$$

Prova.

Pelo teorema da série de Fourier é possível expandir Ψ em um conjunto ortornormal de autofunções de H , ou seja funções que satisfazem a equação

$$H\varphi_k = \epsilon_k \varphi_k, \tag{6-2}$$

pode ser escrito que

$$\Psi = \sum_{k=1}^{\infty} c_k \varphi_k, \tag{6-3}$$

onde

$$c_k = \int \varphi_k^* \psi d\tau, \quad 6-4$$

Como

$$\langle \psi | H | \psi \rangle = \int \psi^* H \psi d\tau, \quad 6-5$$

tem-se

$$\langle \psi | H | \psi \rangle = \int \left(\sum_{k=1}^{\infty} c_k \varphi_k \right)^* H \left(\sum_{k=1}^{\infty} c_k \varphi_k \right) d\tau = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} c_k^* c_l \int \varphi_k^* H \varphi_l d\tau, \quad 6-6$$

Como $H\varphi_k = \varepsilon_k \varphi_k$,

$$\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} c_k^* c_l \varepsilon_k \varepsilon_l \int \varphi_k \varphi_l d\tau, \quad 6-7$$

por definição $\varphi_k \perp \varphi_l$, logo

$$\int \psi^* H \psi d\tau = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} c_k c_l \varepsilon_k \varepsilon_l \delta_{k,l} = \sum_{k=1}^{\infty} c_k^2 \varepsilon_k, \quad 6-8$$

onde δ é a função delta de *kroncker*.

Tomando ε_0 como o menor autovalor de H , $\varepsilon_k \geq \varepsilon_0$, logo,

$$\sum_{k=1}^{\infty} c_k^2 \varepsilon_k \geq \varepsilon_0 \sum_{k=1}^{\infty} c_k^2, \quad 6-9$$

por procedimentos similares $\int \psi^* \psi d\tau = \sum_{k=1}^{\infty} c_k^2$, assim,

$$\frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\sum_{k=1}^{\infty} c_k^2 \varepsilon_k}{\sum_{k=1}^{\infty} c_k^2} \geq \frac{\varepsilon_0 \sum_{k=1}^{\infty} c_k^2}{\sum_{k=1}^{\infty} c_k^2} = \varepsilon_0 \quad 6-10$$

logo,

$$\frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} \geq \varepsilon_0. \blacksquare \quad 6-11$$

Os Algoritmos Evolucionários poderão ser utilizados sempre que houver uma forma de mensurar a qualidade de uma solução proposta. Sendo assim, o mesmo algoritmo pode ser utilizado sem que seja necessária a formulação de novos procedimentos de otimização, e é possível reutilizar o otimizador para resolver problemas diferentes, sendo apenas necessárias as devidas mudanças na função objetivo.

De acordo com o teorema 6.1, achar o estado fundamental de um sistema multieletrônico é equivalente a resolver um problema de otimização cuja função objetivo é o cálculo da energia associada à função de onda.

A aplicação de Algoritmos Evolucionários para resolver problemas de otimização cujo domínio seja real é simples, uma vez que o algoritmo em questão possui representação direta no domínio do problema. Mais especificamente, cada dimensão do espaço de busca é associada a um gene de um indivíduo da população do algoritmo evolutivo.

Para que seja possível encontrar funções de onda evolucionárias é necessário transformar o problema de busca por funções em um problema de otimização numérica. E isso pode ser feito graças à presença de isomorfismos isométricos entre o conjunto de funções aceitáveis como solução para equação de Schrödinger e o conjunto ℓ^2 . Esse isomorfismo pode ser estabelecido, por exemplo, pelas séries de Fourier.

O conjunto das funções de onda trabalhadas nesta tese será o fecho da intersecção entre o conjunto das funções quadrado integráveis, contínuas e de derivadas contínuas até segunda ordem, e simétricas ou antissimétricas com relação à permutação de suas coordenadas. Esse conjunto nesta tese é denominado \mathcal{F}^+ e é ilustrado na Figura 6-2.

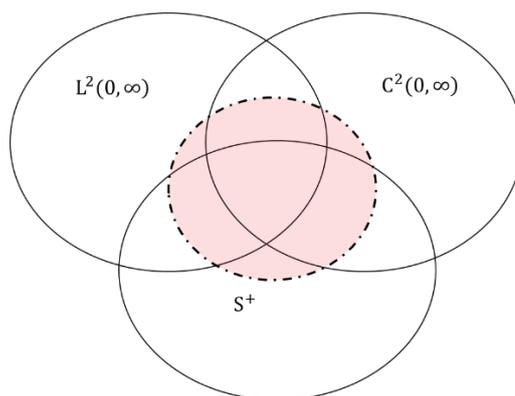


Figura 6-2– O conjunto \mathcal{F}^+

Assim, encontrar a solução da equação de Schrödinger para sistemas eletrônicos é equivalente a encontrar os coeficientes de Fourier da função presente

em \mathcal{F}^+ que corresponde ao estado fundamental do sistema estudado, que é exatamente equivalente a resolver um problema de otimização numérica, uma vez que o funcional de energia é a própria função objetivo, o range numérico do operador de Schrödinger, sendo possível também, definir outras funções objetivos que podem ter os mesmo pontos de mínimo.

Proposição 6.2

Se Ψ é escrito como uma série de Fourier então, a energia total, ϵ_t do sistema é dada com uma soma das contribuições dos elementos φ_i de uma base.

ou seja, $\Psi = \sum_{i=1}^{\infty} c_i \varphi_i$ tem-se

$$\epsilon_t = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{(\sum_{i=1}^{\infty} c_i \varphi_i) H (\sum_{i=1}^{\infty} c_i \varphi_i)}{(\sum_{i=1}^{\infty} c_i \varphi_i) (\sum_{i=1}^{\infty} c_i \varphi_i)} = \frac{\sum_{i,j=1}^{\infty} c_i c_j \int \varphi_i H \varphi_j}{\sum_{i,j=1}^{\infty} c_i c_j \int \varphi_i \varphi_j}, \tag{6-12}$$

$$\epsilon_k = \frac{c_i c_j}{\sum_{i=1}^{\infty} c_i c_j \int \varphi_i \varphi_j} \int \varphi_i H \varphi_j. \blacksquare \tag{6-13}$$

Proposição 6.3

Se

$$\epsilon_t = \sum_k^{\infty} \epsilon_k, \epsilon_k \rightarrow 0.$$

Prova.

$$\sum_k^{\infty} \epsilon_k = \lim_{n \rightarrow \infty} \sum_{k=1}^n \epsilon_k, \tag{6-14}$$

Seja um inteiro $z \geq 1$, tal que a série

$$\lim_{n \rightarrow \infty} \sum_{k=z}^n \epsilon_k < \infty. \tag{6-15}$$

Para provarmos usaremos a seguinte relação válida para todo $z \geq 2$ e $w \geq 2$

$$\sum_{k=1}^{z+w} \epsilon_k - \sum_{k=1}^{z-1} \epsilon_k = \sum_{k=z}^{z+w} \epsilon_k. \tag{6-16}$$

Como $\lim_{n \rightarrow \infty} \sum_{k=1}^n \epsilon_k = \epsilon_t$, como enunciado pela proposição 6.1, então

$$\lim_{n \rightarrow \infty} \sum_{k=z}^n \epsilon_k < \infty, \text{ para qualquer } z \text{ inteiro. Então,}$$

$$\lim_{z \rightarrow \infty} \sum_{k=1}^z \varepsilon_k - \sum_{k=1}^{w-1} \varepsilon_k = \lim_{z \rightarrow \infty} \sum_{k=w}^z \varepsilon_k, \quad 6-17$$

fazendo $w \rightarrow \infty$, tem-se

$$\lim_{z \rightarrow \infty} \sum_{k=1}^z \varepsilon_k - \lim_{w \rightarrow \infty} \sum_{k=1}^{w-1} \varepsilon_k = 0, \quad 6-16$$

portanto,

$$\lim_{z, w \rightarrow \infty} \sum_{k=w}^z \varepsilon_k = 0, \quad 6-17$$

logo, como a serie $\lim_{w \rightarrow \infty} \sum_{k=w}^{\infty} \varepsilon_k$ vai a 0, tem-se que

$$\lim_{k \rightarrow \infty} \varepsilon_k = 0. \blacksquare \quad 6-18$$

Esse conjunto \mathcal{F}^+ deve ter algumas características que possibilitem o uso dos métodos aqui propostos, como por exemplo, todas as funções nele presentes, possam ser descritas com uma série de Fourier. As propriedades que \mathcal{F}^+ tem que possuir serão mostradas e demonstradas a seguir.

Seja $\mathcal{F}^+ = \overline{C^2(0, \infty) \cap L^2(0, \infty) \cap S}$ em $L^2(0, \infty)$, onde C^2 é o conjunto de funções contínuas e de derivadas, até ordem 2, também contínuas, L^2 o conjunto das funções que possuem a integral do seu quadrado finita, e S o conjunto de funções simétricas.

Definição 6.4

Seja $P_{i,j}$ o operador permutação, seja $f: \mathcal{R}^n \rightarrow \mathcal{R}$, Se $f \in S$, então $f(r_1, r_2, \dots, r_n) = f(P_{i,j}(r_1, r_2, \dots, r_n)), \forall i, j \in \mathbb{N}$.

Tomando sistemas de dois elétrons, o átomo de Hélio e seus íons. Exemplos de funções no conjunto \mathcal{F}^+ , são:

$$\Psi'(r_1, r_2) = f(r_1)g(r_2) + f(r_2)g(r_1), \quad 6-19$$

onde:

$$f, g \in L^2(0, \infty) \cap C^2(0, \infty).$$

Proposição 6.5

Sejam as aplicações contínuas $\psi': \mathcal{R}_+^2 \rightarrow \mathcal{R}$, $\psi: \mathcal{R}_+^2 \rightarrow \mathcal{R}$, $f: \mathcal{R}_+ \rightarrow \mathcal{R}$, $g: \mathcal{R}_+ \rightarrow \mathcal{R}$. Onde:

$$\psi(r_1, r_2) = f(r_1)g(r_2),$$

$$\psi'(r_1, r_2) = f(r_1)g(r_2) + f(r_2)g(r_1).$$

Existe ao menos uma f e uma g , ambas em $L^2(0, \infty)$ que satisfaz a seguinte desigualdade.

$$\frac{\langle \psi' | H | \psi' \rangle}{\langle \psi' | \psi' \rangle} < \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle}. \quad 6-20$$

Prova.

Proposição 6.6

O conjunto \mathcal{F}^+ espaço é espaço produto interno e é completo com relação à métrica euclidiana.

Prova

Basta aplicar o teorema do conjunto fechado, todo conjunto fechado dentro de um completo é completo. Assim, tome o fecho de um conjunto arbitrário em $L^2(0, \infty)$, que é completo, com relação a métrica euclidiana, como \mathcal{F}^+ é fechado, então pelo teorema segue que ele é completo.

Proposição 6.7

$$\mathcal{F}^+ \subsetneq L^2.$$

Prova.

$$\mathcal{F}^+ = \overline{C^2(0, \infty) \cap L^2(0, \infty)} \cap \mathcal{S}, \quad 6-21$$

$$\mathcal{F}^+ \subseteq \overline{C^2(0, \infty)} \cap \overline{L^2(0, \infty)} \cap \overline{\mathcal{S}}, \quad 6-22$$

$$\mathcal{F}^+ \subseteq L^2(0, \infty) \cap L^2(0, \infty) \cap \overline{\mathcal{S}}, \quad 6-23$$

$$\mathcal{F}^+ \subseteq L^2(0, \infty) \cap \overline{\mathcal{S}},$$

$$\overline{\mathcal{S}} = \mathcal{S},$$

logo,

$$L^2(0, \infty) \cap \mathcal{S} \subsetneq L^2, \quad 6-24$$

então,

$$\mathcal{F}^+ \subsetneq L^2. \blacksquare \quad 6-25$$

Proposição 6.8

O conjunto \mathcal{F}^+ tem ao menos uma base ortogonal.

Prova.

Todo subconjunto fechado de um espaço de Hilbert é Hilbert, e todo espaço de Hilbert, tem uma base ortogonal contável.

O espaço de busca é grande demais e a aplicação direta do algoritmo evolucionário pode não ser bem sucedida, pois a princípio os coeficientes da série de Fourier que representa a função de onda estão em \mathcal{R}^n , então é importante desenvolver um método para especificação de um subconjunto de \mathcal{R}^n que contenha a resposta procurada. Esse conjunto foi especificado e caracterizado pelo teorema 5.9 e seus corolários. ■

Teorema 6.9

O conjunto de coeficientes de Fourier da expansão da função de onda, Ψ , correspondente ao estado fundamental de qualquer sistema monoelétrônico estará sempre contido no intervalo

$$\left[-\frac{1}{2} \sqrt{\frac{4\zeta + 1}{\pi}}, \frac{1}{2} \sqrt{\frac{4\zeta + 1}{\pi}} \right],$$

desde que o $\sup \Psi^2 \leq \zeta$.

Prova

Tome,

$$r = \sqrt{x^2 + y^2 + z^2}, \tag{6-26}$$

assuma, $\Psi(r)$ normalizado, isto é, tal que

$$\iiint_{-\infty}^{\infty} \Psi(\sqrt{x^2 + y^2 + z^2})^2 dx dy dz = 1, \tag{6-27}$$

$$\begin{aligned} \iiint_{-\infty}^{\infty} \Psi(\sqrt{x^2 + y^2 + z^2})^2 dx dy dz \\ = \left(\int_0^{\infty} \Psi(r)^2 r^2 dr \right) \left(\int_0^{\pi} \sin(\phi) d\phi \right) \left(\int_0^{2\pi} 1 d\theta \right) \\ = 4\pi \int_0^{\infty} \Psi(r)^2 r^2 dr, \end{aligned} \tag{6-28}$$

logo,

$$\int_0^{\infty} \Psi(r)^2 r^2 dr = \frac{1}{4\pi}, \tag{6-29}$$

pelo teorema da série de Fourier, todo ponto em um espaço de Hilbert pode ser escrito como,

$$\psi(r) = \sum_{i=1}^{\infty} c_i \varphi_i(r), \quad 6-30$$

e os coeficientes c_i são dados por

$$c_i = \int_0^{\infty} \psi(r) \varphi_i(r) dr, \quad 6-31$$

aplicando desigualdade de *cauchy-schwartz*, tem-se

$$\left| \int_0^{\infty} \psi(r) \varphi_i(r) dr \right|^2 \leq \left(\int_0^{\infty} \psi(r)^2 dr \right) \left(\int_0^{\infty} \varphi_i(r)^2 dr \right), \quad 6-32$$

assim,

$$|c_i|^2 \leq \left(\int_0^{\infty} \psi(r)^2 dr \right) \left(\int_0^{\infty} \varphi_i(r)^2 dr \right), \quad 6-33$$

dividindo o domínio das integrais em duas partes, $[0,1]$ e $[1,\infty]$

$$\int_0^1 \psi(r)^2 r^2 dr + \int_1^{\infty} \psi(r)^2 r^2 dr = \frac{1}{4\pi}, \quad 6-34$$

$$\int_0^{\infty} \psi(r)^2 dr = \int_0^1 \psi(r)^2 dr + \int_1^{\infty} \psi(r)^2 dr, \quad 6-35$$

é possível estabelecer um limite superior para cada um dos termos da integral, dados por

$$\int_0^1 \psi(r)^2 r^2 dr \leq \frac{1}{4\pi}, \quad 6-36$$

$$\int_1^{\infty} \psi(r)^2 r^2 dr \leq \frac{1}{4\pi}, \quad 6-37$$

nota-se, além disso, as seguintes relações

$$\int_0^1 \psi(r)^2 r^2 dr \leq \int_0^1 \psi(r)^2 dr, \quad 6-38$$

$$\int_1^{\infty} \psi(r)^2 dr \leq \int_1^{\infty} \psi(r)^2 r^2 dr, \quad 6-39$$

$$\int_1^{\infty} \psi(r)^2 dr \leq \int_1^{\infty} \psi(r)^2 r^2 dr \leq \frac{1}{4\pi}, \quad 6-40$$

$$\int_1^{\infty} \psi(r)^2 dr \leq \frac{1}{4\pi}, \quad 6-41$$

tome uma ζ constante tal que,

$$\sup \psi(r)^2 = \lim_{p \rightarrow \infty} \left(\int_0^1 \psi(r)^{2p} dr \right)^{\frac{1}{p}} \leq \zeta, \quad 6-42$$

logo,

$$\int_0^1 \psi(r)^2 dr \leq \int_0^1 \zeta dr \leq \zeta, \quad 6-43$$

$$\int_0^1 \psi(r)^2 dr + \int_1^\infty \psi(r)^2 dr \leq \zeta + \frac{1}{4\pi}, \quad 6-44$$

como φ_i vem de um conjunto ortonormal, então

$$|c_i|^2 \leq \left(\int_0^\infty \psi(r)^2 dr \right), \quad 6-45$$

$$|c_i| \leq \sqrt{\frac{4\zeta + 1}{4\pi}} = \frac{1}{2} \sqrt{\frac{4\zeta + 1}{\pi}}. \blacksquare \quad 6-46$$

Corolário 6.10

Se a função de onda total do sistema é representada por uma função da forma,

$$\Psi(r_1, r_2, \dots, r_n) = \prod_{i=1}^n \varphi_i(r_i), \quad 6-47$$

onde $\varphi_i(r_i)$ é uma função de onda monoelétrica, então, os coeficientes da expansão de Fourier da função de onda que corresponde ao estado fundamental de qualquer sistema multieletrônico obedeceram a seguinte relação:

$$|c_{k,1,\dots}| \leq \frac{1}{2^n} \left| \left(\frac{4C + 1}{\pi} \right)^{\frac{n}{2}} \right|, \quad 6-48$$

onde $|c_{k,1,\dots}|$ são os coeficiente da expansão de Fourier da função de onda estudada.

Desde que C seja uma cota superior para $\Psi(r_1, r_2, \dots, r_n)$.

Prova.

Tome B uma base de Fourier para um espaço de Hilbert de funções. Pelo teorema das series de Fourier, toda função contida nesse espaço tem pelo menos uma expansão do tipo,

$$\varphi_i(r_i) = \sum_{j=1}^{\infty} c_j^i B_j(r_i), \tag{6-49}$$

Faça $n = 2$,

$$\psi(r_1, r_2) = \varphi_1(r_1) \varphi_2(r_2), \tag{6-50}$$

$$\psi(r_1, r_2) = \left(\sum_{j=1}^{\infty} c_j^1 B_j(r_1) \right) \left(\sum_{j=1}^{\infty} c_j^2 B_j(r_2) \right), \tag{6-51}$$

$$\psi(r_1, r_2) = \left(\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} c_l^1 c_k^2 B_k(r_1) B_l(r_2) \right), \tag{6-52}$$

$$\psi(r_1, r_2) = \left(\sum_{k,l=1}^{\infty} c_{k,l} B_k(r_1) B_l(r_2) \right), \tag{6-53}$$

$$|c_{k,l}| = |c_l^1 c_k^2| = |c_l^1| |c_k^2|, \tag{6-54}$$

do teorema (anterior),

$$|c_i| \leq \frac{1}{2} \sqrt{\frac{4\zeta + 1}{\pi}}. \tag{6-55}$$

Logo,

$$|c_{k,l,\dots}| = c_l^1 c_k^2 \dots, \forall k, l \dots \tag{6-56}$$

Por indução,

$$|c_{k,l,\dots}| \leq \prod_{i=1}^n \left| \frac{1}{2} \sqrt{\frac{4\zeta_i + 1}{\pi}} \right| \tag{6-57}$$

onde ζ_i é uma cota superior φ_i para todo i .

Seja a seguinte coleção,

$$\zeta = (\zeta_1, \zeta_2, \dots, \zeta_n)$$

Tome $C = \sup \zeta$,

$$\frac{1}{2^n} \prod_{i=1}^n \left| \sqrt{\frac{4\zeta_i + 1}{\pi}} \right| \leq \frac{1}{2^n} \left| \left(\frac{4C + 1}{\pi} \right)^{\frac{n}{2}} \right| \tag{6-60}$$

Logo,

$$|c_{k,l,\dots}| \leq \frac{1}{2^n} \left| \left(\frac{4C + 1}{\pi} \right)^{\frac{n}{2}} \right| \blacksquare \tag{6-61}$$

Teorema 6.10

Se o número de elétrons no sistema tende a infinito, então o módulo de qualquer dos coeficientes da expansão de Fourier de toda função de onda multieletrônica, que corresponde ao estado fundamental de qualquer sistema, é menor ou igual a 1, desde que o $\sup \varphi_i (r_i) = -\frac{1}{4} + \pi, \forall_i$.

Prova.

Seja

$$|c| \leq \lim_{n \rightarrow \infty} \frac{1}{2^n} \left| \left(\frac{4C + 1}{\pi} \right)^{\frac{n}{2}} \right|, \quad 6-62$$

Fazendo

$$\frac{4C + 1}{\pi} = 4, \quad 6-63$$

Temos

$$|c| \leq \lim_{n \rightarrow \infty} \frac{1}{2^n} |2^n| = 1, \quad 6-64$$

Logo,

$$C = -\frac{1}{4} + \pi. \blacksquare \quad 6-65$$

Teorema 6.10

Se o número de elétrons no sistema tende a infinito então o módulo dos coeficientes da expansão de Fourier de toda função de onda multieletrônica que corresponde ao estado fundamental do sistema tende a zero. Desde o $\sup \varphi_i (r_i) < -\frac{1}{4} + \pi \forall_i$.

Prova.

Seja

$$a = \frac{4C + 1}{\pi}, \quad 6-66$$

$$|c| \leq \lim_{n \rightarrow \infty} \frac{1}{2^n} \left| a^{\frac{n}{2}} \right|, \quad 6-67$$

$$|c| \leq \left(\lim_{n \rightarrow \infty} \frac{1}{2^n} \right) \left(\lim_{n \rightarrow \infty} \left| a^{\frac{n}{2}} \right| \right), \quad 6-68$$

Se $a < 4$ então a função 2^n , presente no denominador varia mais rápido do que $a^{\frac{n}{2}}$. Uma vez que o $\lim_{n \rightarrow \infty} \frac{1}{2^n} = 0$. Então a sequência $\left(\lim_{n \rightarrow \infty} \frac{1}{2^n} \right) \left(\lim_{n \rightarrow \infty} a^{\frac{n}{2}} \right)$ tem limite igual a zero. \blacksquare

Lema. A sequência $\frac{1}{2^n} \left| \left(\frac{4\zeta+1}{\pi} \right)^{\frac{n}{2}} \right|$ é convergente, se e somente se, $\zeta \leq -\frac{1}{4} + \pi$.

Teorema 6.11

Seja um conjunto ortonormal, quando considerada a função peso $W(x)$. Os coeficientes da série de Fourier de qualquer função de onda que corresponda ao estado fundamental de todo sistema monoelétrônico estão contidos no seguinte intervalo

$$\left[-\frac{1}{2} \left(\sqrt{\frac{\alpha_i(4\zeta+1)}{\pi}} \right), \frac{1}{2} \sqrt{\frac{\alpha_i(4\zeta+1)}{\pi}} \right]$$

desde que o $\sup \psi^2 \leq \zeta$. Onde

$$\alpha_i = \int \varphi_i(x)^2 W(x) dx.$$

Prova.

Seja $\psi(r) = \sum_{i=1}^{\infty} c_i \varphi_i(r)$, e os coeficientes c_i são dados por

$$c_i = \int_0^{\infty} \psi(r) \varphi_i(r) w(r)^{\frac{1}{2}} dr, \tag{6-69}$$

aplicando desigualdade de Cauchy-Schwartz, tem-se

$$\left| \int_0^{\infty} \psi(r) \varphi_i(r) w(r)^{\frac{1}{2}} dr \right|^2 \leq \left(\int_0^{\infty} \psi(r)^2 dr \right) \left(\int_0^{\infty} \varphi_i(r)^2 w(r) dr \right), \tag{6-70}$$

assim,

$$|c_i|^2 \leq \left(\int_0^{\infty} \psi(r)^2 dr \right) \left(\int_0^{\infty} \varphi_i(r)^2 w(r) dr \right), \tag{6-71}$$

do teorema 6-1,

$$\int_0^{\infty} \psi(r)^2 dr \leq \zeta + \frac{1}{4\pi}, \tag{6-72}$$

fazendo

$$\int_0^{\infty} \varphi_i(r)^2 w(r) dr = \alpha_i, \tag{6-73}$$

tem-se

$$|c_i|^2 \leq \left(\zeta + \frac{1}{4\pi} \right) \alpha_i, \tag{6-74}$$

logo,

$$|c_i| \leq \left[-\frac{1}{2} \left(\sqrt{\frac{\alpha_i(4\zeta + 1)}{\pi}} \right) \right]. \blacksquare$$

6.1.1.

Breve consideração sobre o funcional de Energia de sistemas multieletrônicos

A energia do estado fundamental do átomo de Hélio tem sido objeto de interesse durante muito tempo e ainda continua atraindo e chamando atenção de muitos pesquisadores. Sempre com o objetivo de extrapolar o conhecimento para sistemas arbitrários. Nesse sentido o método variacional tem sido uma das mais poderosas ferramentas utilizadas para estimar a energia do estado fundamental de qualquer sistema quântico. Muito embora mais a frente será apresentada uma lista aproximações desenvolvidas, ao logo história, para o estado fundamental do Hélio, é importante destacar que, as bases utilizadas para aproximar as soluções da equação de Schrödinger na maioria dos casos da lista, não são necessariamente bases. Isso implica na perda de generalidade dos trabalhos citados a seguir nesta seção. Daí a importância dessa tese, que desenvolve e apresenta métodos para construir funções de base que podem ser utilizadas em funcionais, já desenvolvidos na literatura, tanto o baseado em determinantes de Slater, como o baseado em antissimetização com o operador de Young desenvolvido em (NASCIMENTO, 2007), ou mesmo aqueles que incluem explicitamente o termo r_{12} , como aqueles apresentados a seguir.

6.1.2.

Um Estudo Bibliográfico sobre as funções de onda do estado fundamental do átomo de Hélio

Durante os últimos anos, esse método foi aplicado ao estudo do átomo de Hélio para realizadas as mais diversas tarefas, como exemplo, citam-se: Komasaand & Rychlewski (2003) que considerou funções gaussianas correlacionadas em cálculos variacionais para estimar a energia do estado fundamental do átomo de Hélio, Flores-Riveros *et al* (2003) e Theodorakis *et al.* (2010) consideraram um método de otimização para aproximar a solução de um

conjunto de equações. Lobanova *et al.* (2007) e Biaye *et al.* (2005) usaram o método variacional para o cálculo das energias dos estados excitados do átomo de Hélio para mostrar a importância dos efeitos de correlação eletrônica.

Os cálculos variacionais aplicados ao átomo de Hélio foram sempre considerados como *benchmark* para os testes de novas teorias e é esse o caso mais simples de um sistema real para o qual a equação de Schrödinger ainda não foi resolvida de forma exata.

Conforme já mencionado, a equação de Schrödinger descreve de forma precisa as estruturas eletrônicas de sistemas atômicos e moleculares. A aplicação da equação de Schrödinger ao cálculo de propriedades do átomo de Hidrogênio foi um exemplo da precisão dessa equação. O próximo foi o átomo de Hélio e portanto imediatamente após Schrödinger, Hyrellas (1929) realizou estudos sobre o átomo de Hélio e publicou um resultado muito importante, o qual foi muito preciso, mas ainda longe do resultado exato. Desde então, muitos cientistas tem tentado resolver de forma mais precisa ainda a equação de Schrödinger para o átomo de Hélio, Hylleraas(1929), Kinoshita (1957), (Frankowski & Pekeris,1966), Korobov (2002), (Nakashima & Nakatsujia,2007). Essa solução poderia dar uma ideia sobre a solução dessa equação para sistemas mais gerais. Esta área de pesquisa tem sido chamada de teoria das funções explicitamente correlacionadas.

Hylleraas considerou intuitivamente que as funções de onda do átomo de Hélio poderiam ser expressas com uma boa aproximação como

$$\sum C_{l,m,n} s^l t^m u^n \exp(-\alpha s), \quad 4-10$$

Onde $\{s, t, u\}$ é um sistema de coordenadas definido por

$$s = r_1 + r_2,$$

$$t = r_1 - r_2, \quad 4-11$$

$$u = r_{12},$$

$\{l, m, n\}$ é um conjunto de números inteiros e $C_{l,m,n}$ é um coeficiente a ser determinado. Essa expansão da função de onda foi chamada de *Hylleraas expansion*. Ele introduziu explicitamente a distância inter-eletrônica na função de onda, e isso foi o que resultou num resultado surpreendente quando comparado a teoria de orbitais moleculares. Depois do trabalho pioneiro de Hylleraas,

diferentes funções analíticas incluindo o termo r_{12} foram estudadas das mais diversas formas. A tabela 4.1 apresenta um breve resumo histórico sobre a precisão dos métodos de cálculo do estado fundamental do átomo de Hélio.

Tabela 4.1 – Cronologia do desenvolvimento das melhores funções de onda

Referencia	Tipo da função	Energia (a.u)
1929 Hylleraas	<i>Hylleraas</i>	-2.902 43
1957 Kinoshita	<i>Kinoshita</i>	-2.903 722 5
1966 Frankowski & Pekeris	<i>Logarithm</i>	-2.903 724 377 032 6
1994 Thakkar & Koga	<i>Half-integer</i>	-2.903 724 377 034 114 4
1998 Goldman	<i>Polynomial</i>	-2.903 724 377 034 119 594
1999 Drake	<i>Double exponent</i>	-2.903 724 377 034 119 596
2002 Sims and Hagstrom	<i>12 Hylleraas-CI</i>	-2.903 724 377 034 119 598 29
2002 Korobov	<i>Triple exponent</i>	-2.903 724 377 034 119 598 305
2002 Korobov	<i>Slater geminal</i>	-2.903 724 377 034 119 598 311 158 7
2006 Schwartz	<i>Logarithm lns</i>	-2.903 724 377 034 119 598 311 159 245
2007 Nakashima & Nakatsuji	<i>ICI- Hylleraas</i>	-2.903 724 377 034 119 598 311 159 245 194 404 446 696 905 37

No trabalho original de Hylleraas(1929), o conjunto de inteiros foi limitado aos inteiros não negativos. Kinoshita (1957) verificou que se os inteiros fossem negativos, a precisão dos cálculos poderia ser significativamente melhorada.

Takkar e Koga(1994) estenderam o método Hylleras e obtiveram resultados ainda melhores.

Barllet (1937), Gronwall, (1937) afirmaram que a função de Hylleras não era adequada para descrever a função exata. O argumento dessa afirmação é que a função era suficiente para descrever a condição de *cusp* de duas partículas, mas não descreve a coalescência de três partículas. Então, sugeriram a inclusão de uma singularidade logarítmica na função de onda, e que isso poderia melhorar ainda mais o resultado. Frankowki e Peris(1966) refizeram os cálculos incluindo funções logarítmicas na função de onda e conseguiram funções de onda ainda mais precisas. As suas funções de onda foram expressas como

$$\sum C_{(l,m,n,i,j)} (s^2+t^2)^{\frac{i}{2}} (\ln s)^j \exp(-\alpha s),$$

Onde l, m pertencem aos inteiros, i, j são inteiros não negativos. A forma original

sugerida por Bartlett, Gronwall, e Fock continha o termo $\left(\ln \left[(s^2+t^2)^{\frac{1}{2}}\right]\right)^j$ (xxx), mas o modificaram para um termo mais simples, da forma $\ln s$, isso para tornar o cálculo numérico mais estável. Então, a energia do estado fundamental do átomo de Hélio foi significativamente melhorada.

Os computadores permitiram muitos avanços e cálculos cada vez mais precisos tem sido obtidos. Alguns cálculos utilizando bases ortogonais foram feitos em Goldman(1998). Drade *et al* (1999) realizou cálculos com exponenciais duplas e triplas. Sims & Hagstron (2002) relataram um método chamado de Hyrellas-Configuration com uma grande quantidade de bases. Koborov (2002) realizou cálculos com um número muito grade funções e de coeficientes não lineares e obteve uma energia com uma precisão de 25 casas decimais. Recentemente, usando um conjunto de possíveis bases que incluem termos logaritmos, Schwartz (2006) encontrou uma energia correta com 36 casas decimais. A função apresentada foi ligeiramente diferente das apresentadas em Frankowski e Perikis e foi expressa como

$$\sum C_{(l,m,n,i,j)} s^l (t/s)^m (\ln s)^j \exp(-\alpha s),$$

Onde:

l é um número inteiro;

m, n, i, j são números não negativos.

No artigo *General theorem on Bifurcation and its application to Hartree Equation of the Helium Atom* publicado em 1970 (Reeken,1970) é dada uma prova rigorosa da existência de uma solução positiva da equação de Hartree aplicada ao átomo de Hélio.

Desde que Reeken, em 1970, demonstrou a existência de uma solução positiva para equação radial de Hartree do átomo de Hélio, diversos artigos tem sido escritos sobre as propriedades matemáticas das equações de Schroedinger desde que o trabalho de Reeken foi publicado. Todos esses trabalhos são relacionados com autovalores da equação de Schroedinger. Porém, não são associados a autovalores que são pontos críticos.

Em *Existence and bounds for critical energies of the Hartree operator* (1973), é apresentado um método para provar a existência e calcular *lower bounds* para pontos críticos da equação de *Hartree* aplicada ao átomo de Hélio.

De acordo com Myeres *et al*, 1993, as autofunções de operadores Hamiltonianos de sistemas atômicos e moleculares exibem uma singularidade em todos os pontos no espaço de configurações, onde duas ou mais cargas se encontrem juntas. Na colisão de duas partículas, a derivada da função de onda com relação às coordenadas cartesianas tem uma descontinuidade caracterizada pelas famosas condições de Kato, e a inclusão das condições de Kato nas funções de onda mostram acelerar muito a convergência dos cálculos variacionais de *Rayleigh-Ritz*. Para o caso do Hélio existe também uma singularidade que envolve mais de duas partícula, conhecida como tripla singularidade de colisão quando os três elétrons aproximam-se simultaneamente do núcleo.

6.1.3.

O Funcional de Energia de sistemas de dois elétrons

Nesta seção será desenvolvido o funcional de energia para funções de onda de dois elétrons da forma descrita pela equação 6-76. Mas, todos os cálculos são facilmente estendidos para sistemas de mais elétrons, com auxílio de determinantes de Slater, simetrizadores ou antissimetrizadores de Young, (NASCIMENTO, 2007).

Uma vez que as funções evolucionárias propostas neste trabalho estarão sempre contidas em \mathcal{F}^+ , as funções de onda de dois elétrons, $\psi: \mathcal{R}^2 \rightarrow \mathcal{R}$ poderão ser descritas da seguinte forma:

$$\psi(r_1, r_2) = f(r_1)g(r_2) + f(r_2)g(r_1), \quad 6-76$$

onde:

$$f: \mathcal{R}_+ \rightarrow \mathcal{R},$$

$$g: \mathcal{R}_+ \rightarrow \mathcal{R},$$

r_1 e r_2 são as coordenadas dos elétrons 1 e 2 respectivamente.

$$\frac{\langle \psi' | H | \psi' \rangle}{\langle \psi' | \psi' \rangle} = \frac{\langle f(r_1)g(r_2) + f(r_2)g(r_1) | H | f(r_1)g(r_2) + f(r_2)g(r_1) \rangle}{\langle f(r_1)g(r_2) + f(r_2)g(r_1) | f(r_1)g(r_2) + f(r_2)g(r_1) \rangle} \quad 6-77$$

$$\langle \psi' | H | \psi' \rangle = \langle f(r_1)g(r_2) | H | f(r_1)g(r_2) \rangle + \langle f(r_1)g(r_2) | H | f(r_2)g(r_1) \rangle + \langle f(r_2)g(r_1) | H | f(r_1)g(r_2) \rangle + \langle f(r_2)g(r_1) | H | f(r_2)g(r_1) \rangle \quad 6-78$$

Como H é auto-adjunto, a seguinte relação é sempre válida

$$\langle x | H | y \rangle = \langle y | H | x \rangle. \quad 6-79$$

Sendo x e y ponto arbitrários em um espaço de Hilbert qualquer. Aplicando essa relação na equação (referência) tem-se

$$\langle \psi' | H | \psi' \rangle = \langle f(r_1)g(r_2) | H | f(r_1)g(r_2) \rangle + 2\langle f(r_1)g(r_2) | H | f(r_2)g(r_1) \rangle + \langle f(r_2)g(r_1) | H | f(r_2)g(r_1) \rangle \quad 6-80$$

$$\langle \psi' | \psi' \rangle = \langle f(r_1)g(r_2) | f(r_1)g(r_2) \rangle + 2\langle f(r_1)g(r_2) | f(r_2)g(r_1) \rangle + \langle f(r_2)g(r_1) | f(r_2)g(r_1) \rangle \quad 6-81$$

Por definição o operador H é dado por

$$H = T + V + U \quad 6-82$$

onde,

T é o operador energia cinética;

V é o operador energia potencial elétron-núcleo;

U é operador energia potencial elétron-elétron;

$$\langle \psi | H | \psi \rangle = \langle \psi | T + V + U | \psi \rangle, \quad 6-83$$

$$\langle \psi | H | \psi \rangle = \langle \psi | T | \psi \rangle + \langle \psi | V | \psi \rangle + \langle \psi | U | \psi \rangle, \quad 6-84$$

então,

$$T = -\frac{1}{2} [\nabla_1^2 + \nabla_2^2], \quad 6-85$$

$$\langle \psi | T | \psi \rangle = -\frac{1}{2} \langle \psi | \nabla_1^2 | \psi \rangle - \frac{1}{2} \langle \psi | \nabla_2^2 | \psi \rangle, \quad 6-86$$

em coordenadas esféricas tem-se

$$\nabla_i^2 = \frac{1}{r_i} \nabla_i'^2 = \frac{1}{r_i} \frac{\partial}{\partial r_i} \left(r_i^2 \frac{\partial \psi(r_1, r_2)}{\partial r_i} \right), \quad 6-87$$

normalizando Ψ tem-se,

$$\langle \psi | \nabla_1 | \psi \rangle = 16\pi \int_0^\infty r_2^2 \int_0^\infty -\frac{1}{2} \frac{\psi(r_1, r_2)}{N} \frac{1}{r_i} \frac{\partial}{\partial r_i} \left(r_i^2 \frac{\partial \frac{\psi(r_1, r_2)}{N}}{\partial r_i} \right) r_1^2 dr_1 dr_2,$$

Pela equação (6-76) e por (6-83) tem-se

$$\langle \psi' | T | \psi' \rangle = \langle f(r_1)g(r_2) | T | f(r_1)g(r_2) \rangle + 2\langle f(r_1)g(r_2) | T | f(r_2)g(r_1) \rangle + \langle f(r_2)g(r_1) | T | f(r_2)g(r_1) \rangle, \quad 6-88$$

$$\langle f(r_1)g(r_2) | T | f(r_1)g(r_2) \rangle = \langle f(r_1)g(r_2) | [\nabla_1 + \nabla_2] | f(r_1)g(r_2) \rangle = \langle f(r_1)g(r_2) | \nabla_1 | f(r_1)g(r_2) \rangle + \langle f(r_1)g(r_2) | \nabla_2 | f(r_1)g(r_2) \rangle \quad 6-87$$

$$2\langle f(r_1)g(r_2) | T | f(r_2)g(r_1) \rangle = 2\langle f(r_1)g(r_2) | [\nabla_1 + \nabla_2] | f(r_2)g(r_1) \rangle = 2(\langle f(r_1)g(r_2) | \nabla_1 | f(r_2)g(r_1) \rangle + \langle f(r_1)g(r_2) | \nabla_2 | f(r_2)g(r_1) \rangle) \quad 6-88$$

$$\langle f(r_2)g(r_1) | T | f(r_2)g(r_1) \rangle = \langle f(r_2)g(r_1) | [\nabla_1 + \nabla_2] | f(r_2)g(r_1) \rangle = \langle f(r_2)g(r_1) | \nabla_1 | f(r_2)g(r_1) \rangle + \langle f(r_2)g(r_1) | \nabla_2 | f(r_2)g(r_1) \rangle \quad 6-88$$

$$\langle \psi | \nabla_1 | \psi \rangle = 16\pi \int_0^\infty r_2^2 \int_0^\infty -\frac{1}{2} \frac{\psi(r_1, r_2)}{N} \frac{1}{r_1} \frac{\partial}{\partial r_1} \left(r_1^2 \frac{\partial \frac{\psi(r_1, r_2)}{N}}{\partial r_1} \right) r_1^2 dr_1 dr_2 \quad 6-89$$

$$= \frac{16\pi}{N^2} \left[\int_0^\infty r_2^2 \int_0^\infty -\frac{1}{2} f(r_1)g(r_2) \frac{1}{r_1} \frac{\partial}{\partial r_1} \left(r_1^2 \frac{\partial}{\partial r_1} f(r_1)g(r_2) \right) r_1^2 dr_1 dr_2 + \int_0^\infty r_1^2 \int_0^\infty -\frac{1}{2} f(r_1)g(r_2) \frac{1}{r_2} \frac{\partial}{\partial r_2} \left(r_2^2 \frac{\partial}{\partial r_2} f(r_1)g(r_2) \right) r_2^2 dr_2 dr_1 \right] \quad 6-90$$

$$= -\frac{8\pi}{N^2} \left[\left(\int_0^\infty g(r_2)^2 r_2^2 dr_2 \right) \left(\int_0^\infty f(r_1) \frac{\partial}{\partial r_1} \left(r_1^2 \frac{\partial}{\partial r_1} f(r_1) \right) r_1 dr_1 \right) + \left(\int_0^\infty f(r_1)^2 r_1^2 dr_1 \right) \left(\int_0^\infty g(r_2) \frac{\partial}{\partial r_2} \left(r_2^2 \frac{\partial}{\partial r_2} g(r_2) \right) r_2 dr_2 \right) \right] \quad 6-91$$

$$= \frac{16\pi}{N^2} \left[\int_0^\infty r_2^2 \int_0^\infty -\frac{1}{2} f(r_1)g(r_2) \frac{1}{r_1} \frac{\partial}{\partial r_1} \left(r_1^2 \frac{\partial}{\partial r_1} f(r_2)g(r_1) \right) r_1^2 dr_1 dr_2 + \int_0^\infty r_1^2 \int_0^\infty -\frac{1}{2} f(r_1)g(r_2) \frac{1}{r_2} \frac{\partial}{\partial r_2} \left(r_2^2 \frac{\partial}{\partial r_2} f(r_2)g(r_1) \right) r_2^2 dr_2 dr_1 \right] \quad 6-92$$

$$= -\frac{8\pi}{N^2} \left[\left(\int_0^\infty g(r_2) f(r_2) r_2^2 dr_2 \right) \left(\int_0^\infty f(r_1) \frac{\partial}{\partial r_1} \left(r_1^2 \frac{\partial}{\partial r_1} g(r_1) \right) r_1 dr_1 \right) + \left(\int_0^\infty g(r_1) f(r_1) r_1^2 dr_1 \right) \left(\int_0^\infty g(r_2) \frac{\partial}{\partial r_2} \left(r_2^2 \frac{\partial}{\partial r_2} f(r_2) \right) r_2 dr_2 \right) \right] \quad 6-93$$

$$= \frac{16\pi}{N^2} \left[\int_0^\infty r_2^2 \int_0^\infty -\frac{1}{2} f(r_2) g(r_1) \frac{1}{r_1} \frac{\partial}{\partial r_1} \left(r_1^2 \frac{\partial}{\partial r_1} f(r_2) g(r_1) \right) r_1^2 dr_1 dr_2 + \int_0^\infty r_1^2 \int_0^\infty -\frac{1}{2} f(r_2) g(r_1) \frac{1}{r_2} \frac{\partial}{\partial r_2} \left(r_2^2 \frac{\partial}{\partial r_2} f(r_2) g(r_1) \right) r_2 dr_2 dr_1 \right] \quad 6-94$$

$$= \frac{8\pi}{N^2} \left[\left(\int_0^\infty f(r_2)^2 r_2^2 dr_2 \right) \left(\int_0^\infty g(r_1) \frac{\partial}{\partial r_1} \left(r_1^2 \frac{\partial}{\partial r_1} g(r_1) \right) r_1 dr_1 \right) + \left(\int_0^\infty g(r_1)^2 r_1^2 dr_1 \right) \left(\int_0^\infty f(r_2) \frac{\partial}{\partial r_2} \left(r_2^2 \frac{\partial}{\partial r_2} f(r_2) \right) r_2 dr_2 \right) \right] \quad 6-95$$

Operador energia potencial

Da mesma forma que para operador energia cinética, para o operador potencial tem-se,

$$\langle \psi' | V | \psi' \rangle = \langle f(ar_1)g(\beta r_2) | V | f(ar_1)g(\beta r_2) \rangle + 2 \langle f(ar_1)g(\beta r_2) | V | f(\beta r_1)g(ar_2) \rangle + \langle f(\beta r_1)g(ar_2) | V | f(\beta r_2)g(ar_1) \rangle$$

$$\begin{aligned} \langle f(ar_1)g(\beta r_2) | V | f(ar_1)g(\beta r_2) \rangle &= \langle f(ar_1)g(\beta r_2) | \left[\frac{1}{r_1} + \frac{1}{r_2} \right] | f(ar_1)g(\beta r_2) \rangle \\ &= \langle f(ar_1)g(\beta r_2) | \frac{1}{r_1} | f(ar_1)g(\beta r_2) \rangle \\ &+ \langle f(ar_1)g(\beta r_2) | \frac{1}{r_2} | f(ar_1)g(\beta r_2) \rangle \end{aligned} \quad 6-96$$

$$\begin{aligned} 2 \langle f(ar_1)g(\beta r_2) | V | f(\beta r_1)g(ar_2) \rangle &= 2 \langle f(ar_1)g(\beta r_2) | \left[\frac{1}{r_1} + \frac{1}{r_2} \right] | f(\beta \cdot r_1)g(a \cdot r_2) \rangle \\ &= 2 \langle f(ar_1)g(\beta \cdot r_2) | \frac{1}{r_1} | f(\beta r_1)g(ar_2) \rangle \\ &+ 2 \langle f(ar_1)g(\beta \cdot r_2) | \frac{1}{r_2} | f(\beta r_1)g(ar_2) \rangle \end{aligned} \quad 6-97$$

$$\begin{aligned} \langle f(\beta r_1)g(ar_2) | V | f(\beta r_1)g(ar_2) \rangle &= \langle f(\beta r_1)g(ar_2) | \left[\frac{1}{r_1} + \frac{1}{r_2} \right] | f(\beta r_1)g(ar_2) \rangle \\ &= \langle f(\beta r_1)g(ar_2) | \frac{1}{r_1} | f(\beta r_1)g(ar_2) \rangle \\ &+ \langle f(\beta r_1)g(ar_2) | \frac{1}{r_2} | f(\beta r_1)g(ar_2) \rangle \end{aligned} \quad 6-98$$

$$5 - 96 = \frac{16\pi}{N^2} \left[\int_0^\infty r_2^2 \int_0^\infty -\frac{1}{2} f(ar_1)g(\beta r_2) \frac{1}{r_1} f(ar_1)g(\beta r_2)r_1^2 dr_1 dr_2 + \int_0^\infty r_1^2 \int_0^\infty -\frac{1}{2} f(ar_1)g(\beta r_2) \frac{1}{r_2} f(ar_1)g(\beta r_2)r_2^2 dr_1 dr_2 \right] \quad 6-99$$

$$eq8 = \frac{16\pi}{N^2} \left[\int_0^\infty g(\beta r_2)^2 r_2^2 \int_0^\infty -\frac{1}{2} f(ar_1)^2 r_1 dr_1 dr_2 + \int_0^\infty f(ar_1)^2 r_1^2 \int_0^\infty -\frac{1}{2} g(\beta r_2)^2 r_2 dr_1 dr_2 \right] \quad 6-100$$

$$= 2 \frac{16\pi}{N^2} \left[\int_0^\infty r_2^2 \int_0^\infty -\frac{1}{2} f(ar_1)g(\beta r_2) \frac{1}{r_1} f(\beta r_1)g(ar_2)r_1^2 dr_1 dr_2 + \int_0^\infty r_1^2 \int_0^\infty -\frac{1}{2} f(ar_1)g(\beta r_2) \frac{1}{r_2} f(\beta r_1)g(ar_2)r_2^2 dr_1 dr_2 \right] \quad 6-101$$

$$= 2 \frac{16\pi}{N^2} \left[\int_0^\infty r_2^2 \int_0^\infty -\frac{1}{2} f(ar_1)g(\beta r_2)f(\beta r_1)g(ar_2) r_1 dr_1 dr_2 + \int_0^\infty r_1^2 \int_0^\infty -\frac{1}{2} f(ar_1)g(\beta r_2)f(\beta r_1)g(ar_2) r_2 dr_1 dr_2 \right] \quad 6-102$$

$$= 2 \frac{16\pi}{N^2} \left[\int_0^\infty g(\beta r_2)g(ar_2)r_2^2 \int_0^\infty -\frac{1}{2} f(ar_1)f(\beta r_1) r_1 dr_1 dr_2 + \int_0^\infty f(ar_1)f(\beta r_1)r_1^2 \int_0^\infty -\frac{1}{2} g(\beta r_2)g(ar_2) r_2 dr_1 dr_2 \right] \quad 6-103$$

$$5 - 97 = \frac{16\pi}{N^2} \left[\int_0^\infty r_2^2 \int_0^\infty -\frac{1}{2} f(\beta r_1)g(ar_2) \frac{1}{r_1} f(\beta r_1)g(ar_2)r_1^2 dr_1 dr_2 + \int_0^\infty r_1^2 \int_0^\infty -\frac{1}{2} f(\beta r_1)g(ar_2) \frac{1}{r_2} f(\beta r_1)g(ar_2)r_2^2 dr_1 dr_2 \right] \quad 6-104$$

$$eq10 = \frac{16\pi}{N^2} \left[\int_0^\infty g(ar_2)^2 r_2^2 \int_0^\infty -\frac{1}{2} f(\beta r_1)^2 r_1 dr_1 dr_2 + \int_0^\infty f(\beta r_1)r_1^2 \int_0^\infty -\frac{1}{2} g(ar_2)^2 r_1 dr_1 dr_2 \right] \quad 6-105$$

$$\langle \psi' | U | \psi' \rangle = \langle f(ar_1)g(\beta r_2) | \frac{1}{|r_1 - r_2|} | f(ar_1)g(\beta r_2) \rangle + 2 \langle f(ar_1)g(\beta r_2) | \frac{1}{|r_1 - r_2|} | f(\beta r_1)g(ar_2) \rangle + \langle f(\beta r_1)g(ar_2) | \frac{1}{|r_1 - r_2|} | f(\beta r_2)g(ar_1) \rangle \quad 6-106$$

$$5 - 98 = \frac{16\pi}{N^2} \left[\int_0^\infty r_2^2 \int_0^\infty f(ar_1)g(\beta r_2) \frac{1}{|r_1 - r_2|} f(ar_1)g(\beta r_2)r_1^2 dr_1 dr_2 + 2 \int_0^\infty r_1^2 \int_0^\infty f(ar_1)g(\beta r_2) \frac{1}{|r_1 - r_2|} f(\beta r_1)g(ar_2)r_2^2 dr_1 dr_2 + \int_0^\infty r_1^2 \int_0^\infty f(\beta r_1)g(ar_2) \frac{1}{|r_1 - r_2|} f(\beta r_1)g(ar_2)r_2^2 f dr_1 dr_2 \right] \quad 6-107$$

Para calcular o termo $\langle \psi | U | \psi \rangle$, onde U é a energia de interação elétron-elétron, e é definido por

$$\frac{1}{|r_1 - r_2|} \quad 6-108$$

num sistema com dois elétrons, onde o potencial de interação é colombiano. Apesar de existirem diversos métodos para realizar esse cálculo, neste trabalho optou-se por utilizar transformadas de Fourier.

Para realizar esse cálculo, toma-se o potencial de Coulomb com um caso particular do potencial de Yukawa, definido pela seguinte equação(6-109) quando $u \rightarrow 0$ A transformada de Fourier do potencial de *Coulomb* não converge diretamente, por isso, realiza-se esse artifício.

$$U(x) = -g^2 \frac{e^{-ux}}{x}, \quad 6-109$$

fazendo $g = 1$,

$$\int d^3x \frac{1}{|x|} e^{-iqx} = \lim_{u \rightarrow 0} \int d^3x \frac{e^{-u|x|}}{|x|} e^{-iqx}, \quad 6-110$$

$$\lim_{u \rightarrow 0} 2\pi \int dr r^2 \int d\vartheta \cos(\vartheta) e^{-u|r|} e^{-iqr \cos(\vartheta)}, \quad 6-111$$

$$2\pi \lim_{u \rightarrow 0} \int dr \frac{r}{-iq} e^{-ur - iqr \cos(\vartheta)}, \quad 6-112$$

$$2\pi \lim_{u \rightarrow 0} \int dr \frac{r}{-iq} [e^{-(u+iq)r} - e^{-(u-iq)r}], \quad 6-113$$

$$\frac{2\pi}{-iq} \lim_{u \rightarrow 0} \int dr [r e^{-(u+iq)r} - r e^{-(u-iq)r}], \quad 6-114$$

Lembrando que

$$\int_0^\infty e^{-ar} = -\frac{1}{a} e^{-ar} \Big|_0^\infty = \frac{1}{a}. \quad 6-115$$

Logo,

$$-\frac{2\pi}{iq} \lim_{u \rightarrow 0} \left[\frac{1}{(u+iq)r} - \frac{1}{(u-iq)r} \right], \quad 6-116$$

$$-\frac{2\pi}{iq} \lim_{u \rightarrow 0} \left[\frac{-2iq}{(u+q^2)} \right] = \frac{4\pi}{q^2}. \quad 6-117$$

Então,

$$I = \langle \psi | U | \psi \rangle = \langle \psi | \frac{1}{|r_1 - r_2|} | \psi \rangle, \quad 6-118$$

$$\frac{1}{|r_1 - r_2|} = \int \frac{dk^3}{(2\pi)^3} \frac{4\pi q^2}{k^2} e^{-ik(|r_1 - r_2|)}, \quad 6-120$$

se $\psi(r_1, r_1) = \varphi(r_1)\varphi(r_2)$,

$$I = \iint \varphi(r_1)\varphi(r_2) \frac{1}{|r_1 - r_2|} \varphi(r_1)\varphi(r_2) d^3r_1 d^3r_2, \quad 6-121$$

$$I = \int \frac{d^3k}{(2\pi)^3} \frac{4\pi q^2}{k^2} \int d^3r_1 \varphi(r_1)^2 e^{-ik(|r_1 - r_2|)} \int d^3r_2 \varphi(r_2)^2 e^{-ik(|r_1 - r_2|)}. \quad 6-122$$

Notar que a transformada de Fourier de Φ acontece duas vezes. Usando coordenadas esféricas. Nota-se a presença do conhecido elemento sólido

$$\int d\Omega e^{-ik(|r_i|)} \int d\Omega e^{-ik(|r_i|)} = \int_0^\pi \int_0^{2\pi} \text{sen}(\theta) e^{-ikr} d\theta = 4\pi \frac{\text{sen}(kr)}{kr}, \quad 6-123$$

$$\int d^3r_1 \varphi(r_1)^2 e^{-ik(|r_1 - r_2|)} = 4\pi \int_0^\infty r^2 \varphi(r_1)^2 \frac{\text{sen}(kr)}{kr} dr, \quad 6-124$$

$$\begin{aligned} & \int \frac{d^3k}{(2\pi)^3} \frac{4\pi q^2}{k^2} \int d^3r_1 \varphi(r_1)^2 e^{-ik(|r_1 - r_2|)} \\ &= \int \frac{d^3k}{(2\pi)^3} \frac{4\pi q^2}{k^2} \int_0^\infty r_1^2 \varphi(r_1)^2 4\pi \frac{\text{sen}(kr_1)}{kr_1} dr_1 \int_0^\infty r_2^2 \varphi(r_2)^2 4\pi \frac{\text{sen}(kr_2)}{kr_2} r_2, \end{aligned} \quad 6-125$$

aplicando o (teorema de *Fubini* ver KUBRUSLY,2006) e integrando em k ,

$$32\pi \int_0^\infty r_1^2 \varphi(r_1)^2 \int_0^\infty r_2^2 \varphi(r_2)^2 r_2 \int_0^\infty \frac{1}{k^2} \frac{\text{sen}(kr_1)}{kr_1} \frac{\text{sen}(kr_2)}{kr_2} dk dr_1 dr_2, \quad 6-126$$

$$\int_0^\infty \frac{\text{sen}(kr_1)}{kr_1} \frac{\text{sen}(kr_2)}{kr_2} dk = -\frac{1}{4}\pi((r_1 - r_2)\text{sign}(r_1 - r_2) - (r_1 + r_2)), \quad 6-127$$

$$32\pi \int_0^\infty r_1^2 \varphi(r_1)^2 \int_0^\infty r_2^2 \varphi(r_2)^2 r_2 \int \frac{1}{k^2} \frac{\text{sen}(kr_1)}{kr_1} \frac{\text{sen}(kr_2)}{kr_2} dk dr_1 dr_2, \quad 6-128$$

$$32\pi \int_0^\infty r_1^2 \varphi(r_1)^2 \int_0^\infty r_2^2 \varphi(r_2)^2 \left(-\frac{1}{4}\pi((r_1 - r_2)\text{sign}(r_1 - r_2) - (r_1 + r_2)) \right) dr_1 dr_2, \quad 6-129$$

assim,

$$\begin{aligned} & \langle \psi \left| \frac{1}{|r_1 - r_2|} \right| \psi \rangle \\ &= 8\pi^2 \int_0^\infty r_1^2 \varphi(r_1)^2 \int_0^\infty r_2^2 \varphi(r_2)^2 \left(-\frac{1}{4}\pi((r_1 - r_2)\text{sign}(r_1 - r_2) - (r_1 + r_2)) \right) dr_1 dr_2. \end{aligned} \quad 6-130$$

6.2.

Solução da equação de Schrödinger por algoritmos co-evolucionários com inspiração quântica

Alguns problemas de programação não-linear apresentam características adicionais, tais como, dependência entre as variáveis, e isso pode tornar o problema em questão mais complicado de ser resolvido. No caso em questão serão otimizados dois conjuntos de parâmetros, os parâmetros que caracterizam a base e os coeficientes da série de Fourier utilizada como expansão da função de onda, onde a função objetivo é dada pela equação 6-1.

Existem duas razões principais pelas quais Algoritmos Evolucionários convencionais não são totalmente adequados para resolver este tipo de problema. Em primeiro lugar, os AGs convencionais impedem, em longo prazo, a preservação de certos componentes da solução, pois, por estarem codificados por completo em um indivíduo, eles são avaliados como um todo e apenas os subcomponentes que pertencem a indivíduos com avaliações altas serão preservados. Em segundo lugar, o fato da representação estar relacionada a uma solução completa e por não haver interações entre os membros da população, não existe pressão evolucionária para a ocorrência de co-adaptação, ou seja, não existe pressão para a adaptação de um subcomponente dada à ocorrência de uma mudança em outro subcomponente (*CRUZ et al, 2006*).

No modelo co-evolucionário, duas ou mais espécies formam um ecossistema. Como na natureza, as espécies são geneticamente isoladas, isto é, cada indivíduo pode somente se reproduzir juntamente com outro indivíduo da mesma espécie. Isto é obtido simplesmente isolando cada espécie em uma população separada. As diferentes espécies somente interagem umas com as outras através de um domínio compartilhado e possuem uma relação de cooperação.

A utilização de computação evolucionária possibilita mais do que a otimização de parâmetros, pois, o algoritmo otimizador pode resolver com facilidade problemas cujo domínio seja um espaço inteiro e um contínuo ao mesmo tempo. Dessa forma, é possível especificar o número de bases, que irão ser combinadas para representar cada função de onda evolutiva. Essa característica também torna possível a redução de bases nas funções representadas. Para

resolver esse problema, classificado como problema de programação mista³, podem ser aplicados algoritmos co-evolucionários, pois além de especificar o número de termos da combinação linear, também conseguem boas parametrizações para os coeficientes dos termos utilizados para representar cada uma das funções representadas.

Nesse trabalho é empregado o conceito de co-evolução cooperativa, inspirada na definição de simbiose⁴. Na co-evolução cooperativa, duas ou mais espécies interagem e colaboram para a evolução uma da outra através de um modelo domínio⁵. Nesse modelo, para que as soluções possam ser devidamente avaliadas, essas são montadas a partir da participação de todas as espécies envolvidas. A montagem da solução é feita concatenando os cromossomos de cada espécie, como é mostrado na Figura 6-3.

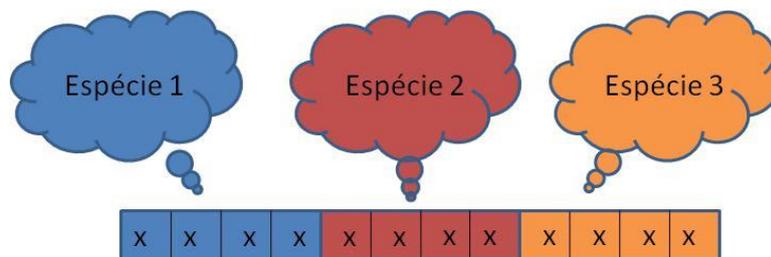


Figura 6-3 Colaboração de vários indivíduos compondo um cromossomo.

Apesar de serem mostradas apenas três espécies neste modelo, o mesmo pode ser usado para n espécies diferentes. Cada espécie evolui em sua própria população e se adapta ao ambiente através de repetidas aplicações do algoritmo evolucionário. Mas, o modelo proposto utiliza esquema de evolução proposto em (CRUZ *et al*, 2005) para cada espécie, conforme discutido detalhadamente no capítulo 3.

A interação entre as espécies no modelo cooperativo é feita seguindo a seguinte sequência de passos:

- 1) Seleciona-se o melhor indivíduo da população i ;

³ Denomina-se programação mista um problema de otimização cuja função custo possui variáveis contínuas e discretas.

⁴ Simbiose é uma relação mutuamente vantajosa entre dois ou mais organismos vivos de espécies diferentes.

⁵ Um modelo domínio é um mecanismo que constrói uma solução a partir da junção de vários cromossomos, segundo um conjunto de regras pré-estabelecidas.

- 2) Formam-se conjuntos de indivíduos concatenando todos os indivíduos, colaboradores, das demais populações como ilustrado pela Figura 6-4;
- 3) Posteriormente, o indivíduo selecionado no passo um é concatenado com todos os indivíduos formados no passo dois;
- 4) As soluções são avaliadas e as aptidões são atribuídas aos indivíduos construídos no passo dois;
- 5) O processo é iterado e executado até todos os indivíduos terem sido avaliados.

O modelo co-evolucionário cooperativo genérico com inspiração quântica é ilustrado na Figura 6-4.

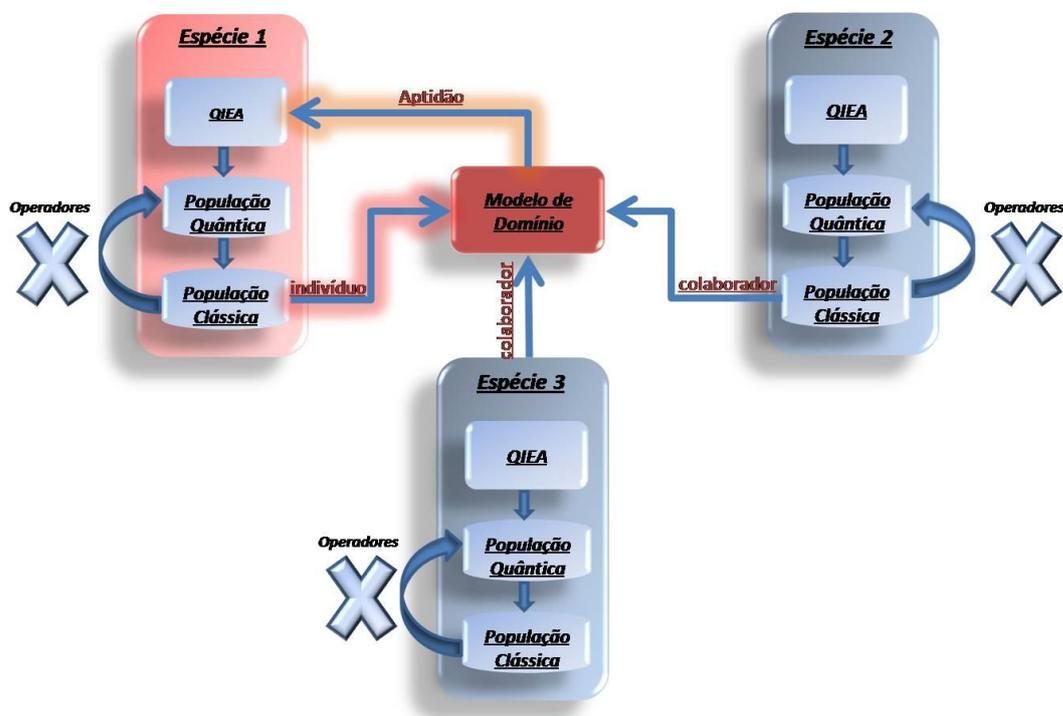


Figura 6-4 Esquema de co-evolução empregando três espécies.

A dinâmica de um processo co-evolutivo conduz a pontos com características especiais. Em (BUCCI, 2007) é mostrado que esses pontos podem ser caracterizados como equilíbrios de *Nash*.

A construção completa de uma função de onda evolucionária é feita realizando-se duas tarefas: a determinação dos coeficientes lineares e não-lineares do modelo de aproximação. Como essas duas tarefas são diretamente relacionadas, cabe a utilização de algoritmos co-evolucionários.

6.2.

Cromossomas propostos para as evoluções que consideram utilização de funções de base

6.2.1.

Cromossoma - para funções de base construídas a partir de polinômios de Laguerre

Para esse tipo de funções o cromossoma é composto apenas de genes reais, onde todos os parâmetros da função aproximada são mapeados em genes reais, tal como na Figura 6-5

$$\alpha_1^\varphi \cdots \alpha_n^\varphi$$

Figura 6-5 Cromossoma de genes reais

Nesse caso as funções consideradas serão dadas por

$$\psi(x) = \sum_{n=1}^{\infty} \alpha_n q_n(x, a)$$

6.2.2.

Cromossoma do tipo 1- para funções de base construídas a partir de wavelets

Propõe-se inicialmente um cromossoma simples, onde todos os parâmetros da função aproximada são mapeados em genes reais, tal como na Figura 6-6

$$\alpha_1^\varphi \cdots \alpha_n^\varphi \quad \alpha_{1,n}^\psi \cdots \alpha_{j,m}^\psi$$

Figura 6-6 Cromossoma de genes reais

Um algoritmo evolutivo que utiliza essa abordagem apenas fornece um conjunto de parâmetros ótimos para funções definidas na equação 6-131. Mas, as funções de onda desse tipo tem pelo menos $n(1 + m)$ parâmetros, e não necessariamente são construídas pelas bases mais importantes em sua composição.

6.2.2.1.

Decodificação do cromossoma do tipo 2- para funções de base construídas a partir de wavelets

Após o processo de decodificação a função utilizada como tentativa para minimização do funcional de energia será dada pela equação

$$\psi(x) = \sum_{k=0}^n \alpha_k^\varphi \overline{\Phi}_k(x - k, K_0, K_1) + \sum_{i=0}^n \sum_{j=0}^m \alpha_{i,j}^\psi \left[2^{\frac{i}{2}} \overline{\Psi}_k(2^i x - j, K_0, K_1) \right] \quad 6-131$$

onde,

$\alpha_1^\varphi \dots \alpha_n^\varphi$ e $\alpha_{1,n}^\psi \dots \alpha_{j,m}^\psi$ corresponde ao vetor de coeficientes da combinação linear utilizadas na aproximação, e pertencem aos reais;

6.2.3.

Cromossoma do tipo 2

Para realizar a otimização do funcional de energia desenvolvido a partir do princípio variacional da equação de Schrödinger aplicando algoritmos evolutivos. Sugere-se um cromossomo composto de partes da forma descrita na Figura 6-7.

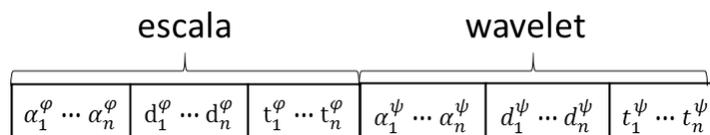


Figura 6-7 Cromossoma sugerido para a representação das funções de onda de cada elétron

Onde,

$\alpha_1^\varphi \dots \alpha_n^\varphi$ ou $\alpha_1^\psi \dots \alpha_n^\psi$ corresponde ao vetor de coeficientes da combinação linear utilizadas na aproximação, e pertencem aos reais;

$d_1^\varphi \dots d_n^\varphi$ ou $d_1^\psi \dots d_n^\psi$ corresponde ao vetor de operadores de dilatação, pertencem ao conjunto dos inteiros;

$t_1^\varphi \dots t_n^\varphi$ ou $t_1^\psi \dots t_n^\psi$ corresponde ao vetor de operadores de dilatação. Nesse caso

t_n^φ pertence ao conjunto dos inteiros;

Com essa abordagem serão especificadas pelo o algoritmo as bases que mais contribuem para a diminuição da energia do sistema estudado e consequentemente com a aproximação da função aproximada.

6.2.3.1.

Decodificação do cromossoma do tipo 2

Para cada cromossoma da Figura 6-7 o processo de decodificação basicamente consiste em posicionar cada um dos genes em posições determinadas segundo conjunto de equações acopladas 6-132, dando origem à função de onda a ser utilizada no método variacional.

$$\left\{ \begin{array}{l} \psi(x) = \sum_{k=1}^n \alpha_k^\varphi \overline{\Phi}_k(x, K_0, K_1) + \sum_{k=1}^m \alpha_k^\psi \overline{\Psi}_k(x, K_0, K_1) \\ \overline{\Phi}_k(x) = 2^{\frac{d_k^\varphi}{2}} \phi(2^{d_k^\varphi} x - t_k^\varphi) \\ \overline{\Psi}_k(x) = 2^{\frac{d_k^\psi}{2}} \phi(2^{d_k^\psi} x - t_k^\psi) \end{array} \right. \quad 6-132$$

ESTUDOS DE CASOS E DISCUSSÃO DE RESULTADOS

Neste capítulo são apresentadas algumas funções de base para cálculos de estrutura que foram construídas de acordo com as técnicas propostas. A função objetivo a ser minimizada é a energia correspondente a função de onda de cada sistema. Os cálculos desenvolvidos nesta tese foram implementados e desenvolvidos com o auxílio do Matlab 2013 e do Visual Studio 2010.

Aqui são realizados experimentos preliminares que objetivam definir a metodologia a ser aplicada posteriormente na construção de funções de base para os átomos B, C, N, O, F, Ne.

Para que fosse possível realizar o estudo dos diversos casos propostos foi necessário implementar os AEIQs e a sua extensão co-evolucionária com os detalhes já mencionados anteriormente. As implementações desses algoritmos foram realizadas em C#. Os estudos de casos presentes neste capítulo são:

1. Análise de desempenho de AEIQs e AEIQs Co-evolutivos aplicados à construção de funções de base;
2. Aplicação de AEIQs à construção de funções de base para o átomo de Lítio;
3. Aplicações do modelo co-evolucionário com inspiração quântica à construção de funções de base;
4. Algoritmos Co-evolutivos com inspiração Quântica e a ortogonalização de Gram–Schmidt;
5. Algoritmos Co-evolutivos com inspiração Quântica e a ortogonalização de Gram–Schmidt paramétrico;

A seguir serão mostrados os estudos de casos que conduziram a definição da metodologia final para a geração de novas das novas funções de base, criadas a partir dos métodos propostos.

7.1.**Estudo de caso 1 - Análise do desempenho de AEIQs e AEIQs Co-evolutivos aplicados à construção de funções de base**

Nesse estudo de caso são empregados os AEIQs e sua extensão co-evolutiva para sintetizar uma função de base para o átomo de Lítio.

A síntese de uma nova função de base para qualquer átomo parte da minimização do funcional de energia associado a cada função de onda. Esse funcional é construído a partir do formato da função de onda. Os experimentos são realizados partir dos funcionais derivados de determinantes de Slater.

Esse funcional resultara na energia de Hartree Fock de uma átomo. A minimização desse funcional deve ser realizada sobre a restrição de ortonormalidade dos spin-orbitais do determinante de Slater. Isso pode ser conseguido por métodos diretos ou utilizando funções de penalização. O funcional associado a energia de determinantes de Slater é a equação 7-1

$$E_{hf} = \sum_{i=1}^n \langle \psi_i, h_i \psi_i \rangle + \sum_{i=1}^n \sum_{j=i+1}^n (J_{ij} - K_{ij}) \quad 7-1$$

Onde :

h_i é a energia de ordem zero do i-ésimo elétron;

J_{ij} é resultado da aplicação do operador de coulomb;

K_{ij} é resultado da aplicação do operador de troca;

É importante lembrar, tal como descrito no capítulo 2, que o operador de Coulomb é responsável pela energia associada a repulsão entre cargas, já o operador de troca atua apenas entre os spin orbitais de mesmo spin (PAGILIA, 1971).

Neste estudo de caso foi escolhida a configuração eletrônica do Li é $1s^2 2s$. Dessa forma, duas aproximações são feitas em paralelo, uma para o formato do orbital 1s e outra para o formato do orbital 2s.

A função tentativa será do seguinte formato $\langle \det 1s\alpha(1) 1s\beta(1) 2s\beta(1) \rangle$. Assim, o funcional a ser minimizado contemplará os seguintes termos

$$\sqrt{[\hat{E}_{hf} - (E_{\text{LimiteHf}} Li) -]^2} - \sqrt{\langle \psi_j, \psi_j \rangle^2} \quad 7-2$$

Nesse caso o E_{LimiteHf} é -7.432726927.

Nesse caso existem dois elétrons $1s$ e apenas 1 do tipo $2s$. Assim, efetivamente a expressão para o cômputo da energia pode ser simplificada é

$$E_{hf} = 2\langle 1s h_i 1s \rangle + \langle 2s h_i 2s \rangle + J_{1s1s} + 2J_{1s2s} + J_{2s2s} - K_{1s2s} \quad 7-3$$

A introdução do termo $\sqrt{\langle \psi_j, \psi_j \rangle^2}$ é responsável por impor a condição de ortogonalidade a minimização.

Na figura são mostradas as curvas de evolução de três rodadas do AEIQ parametrizado conforme a

Tabela 7-1.

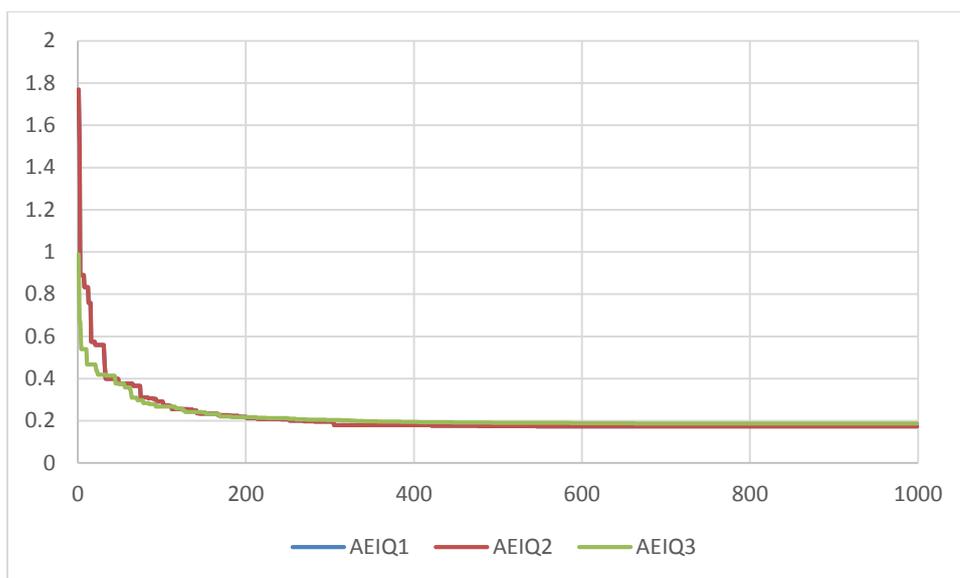


Figura 7-1 Análise de convergência do AEIQ aplicado ao átomo de Li 1s22s (caso1)

Tabela 7-1 Múltiplas configurações de algoritmos evolutivos.

Parâmetro	<i>AEIQ</i> ¹	<i>AEIQ</i> ²	<i>AEIQ</i> ³
Gerações	1000	1000	1000
Tamanho da população clássica	100	100	100
Tamanho da população quântica	2	2	3
Número de observações	2	5	5
Cruzamento Quântico	0.95	0.95	0.95
Cruzamento Clássico	0.95	0.95	0.95
Atualização a cada	10 g	10 g	10 g
Variação de largura	0.1	0.1	0.1
Gap	4	10	14

Observando as curvas de evolução expostas na figura 7-1 que o algoritmo converge 0.2

Apesar do resíduo final do final atingido pelo ao algoritmo estar em torno de 0.2, o que representa um valor alto, é possível perceber que parte desse valor é composto em função do produto interno entre os orbitais não ser zero, ou seja, os orbitais finais não são ortornormais. Esse fato indica que, apesar da convergência do algoritmo, é necessária a adição de mais funções para a construção dos spins-orbitais, ou realizar a aproximação em escalas diferentes, com a introdução de um fator multiplicador no argumento da aproximação.

7.2.

Estudo de caso 2

7.2.1.

Caso 2 a - Aplicação de AEIQs à construção de funções de base para o átomo de Lítio

Nesse estudo de caso será também realizada uma aproximação para os spins orbitais do átomo de Li. A Função objetivo será a mesma do estudo de caso 1, tal como a configuração eletrônica. Mas, serão introduzidos dois parâmetros multiplicadores nos argumentos, esses parâmetros serão os responsáveis por contrair ou dilatar os spin-orbitais. Tais parâmetros serão responsáveis por acelerar a convergência do algoritmo. A fim de verificar o aumento da velocidade de convergência as curvas de evolução mostradas na Figura 7-2.

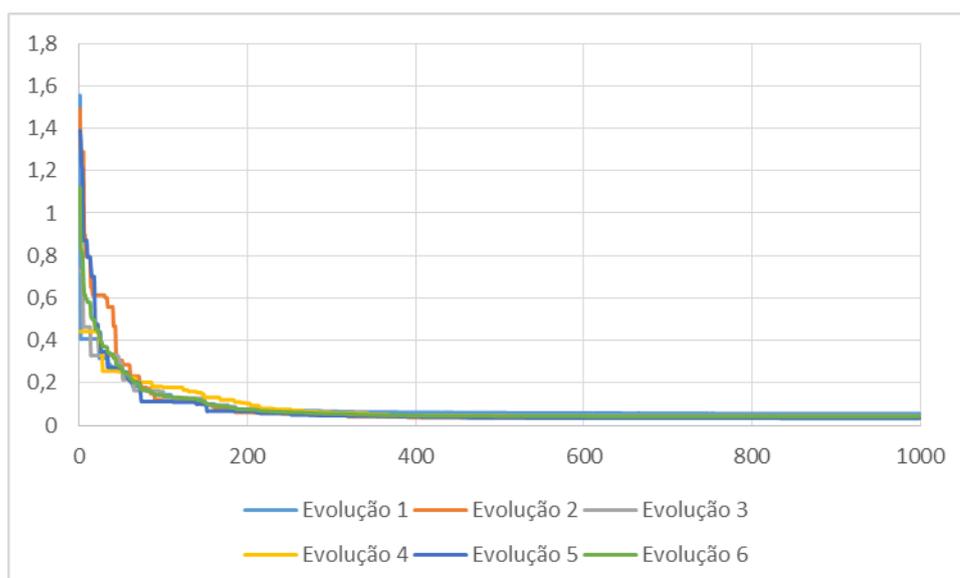


Figura 7-2 Análise de convergência do AEIQ aplicado ao átomo de Li 1s22s (caso2)

A energia final obtida ao término da execução de todos os seis experimentos foi de -7.41504. O cromossomo correspondente é apresentado na Tabela 7-2.

Tabela 7-2 função de base com V=2 e W=2 para o Li 1s2s

g1	g2	g3	g4	g5	g6
-2.1469	-0.1963	-0.0134	0.0000	-0.0030	-1.1229
g7	g8	g9	g10	g11	G12
-0.0835	0.0003	-0.0085	-0.0031	-0.0082	-0.1467
g13	g14	g15	g16	g17	G18
-0.0402	0.0010	0.0085	-0.0100	-0.0015	0.0000
g19	g20	g21	g22	g23	G24
-0.0088	2.6274	-0.0843	-0.0027	-0.0791	-0.0016
g25	g26	g27	g28	g29	G30
0.4812	0.2872	-0.0002	-0.0130	0.0501	-1.2128
g31	g32	g33	g34	g35	G36
-0.0031	-0.0105	-0.0579	0.0168	0.0153	0.0220
g37	g38	g39	g40		
-0.0022	-0.0147	1.1268	0.4128		

As Figura 7-3 e Figura 7-4 mostram as curvas referentes à decodificação do melhor indivíduo, obtido ao término do conjunto de experimentos. As curvas representam a aproximação da função referente ao orbital, sua primeira derivada e a segunda derivada. Essas funções são obtidas naturalmente pela utilização dos procedimentos propostos nesta tese. A obtenção natural dessas funções simplifica

a realização dos cálculos necessários durante o cômputo da energia do sistema estudado.

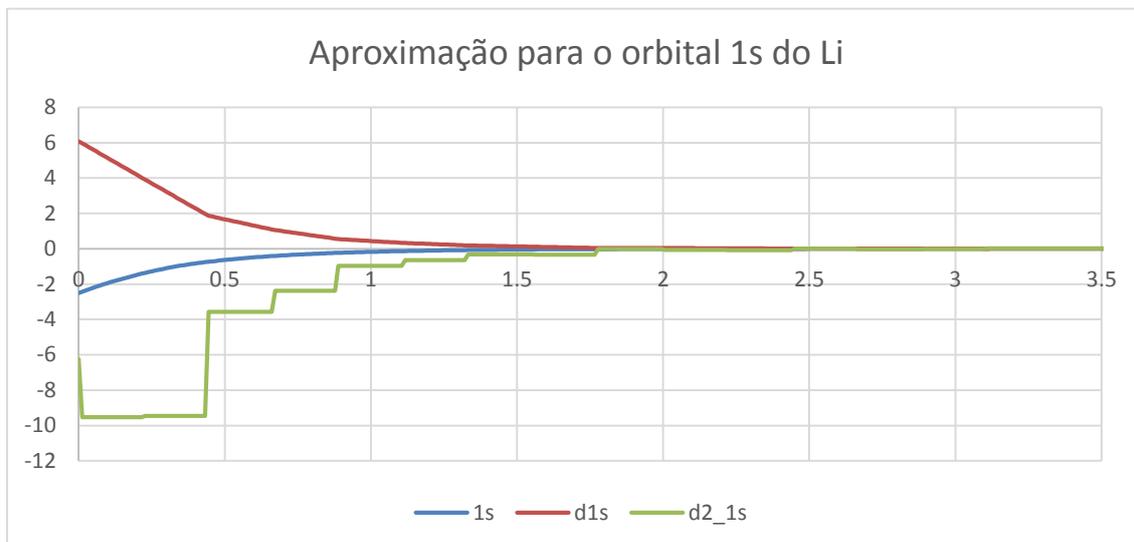


Figura 7-3 a aproximação do orbital 1s do Li, e suas derivadas

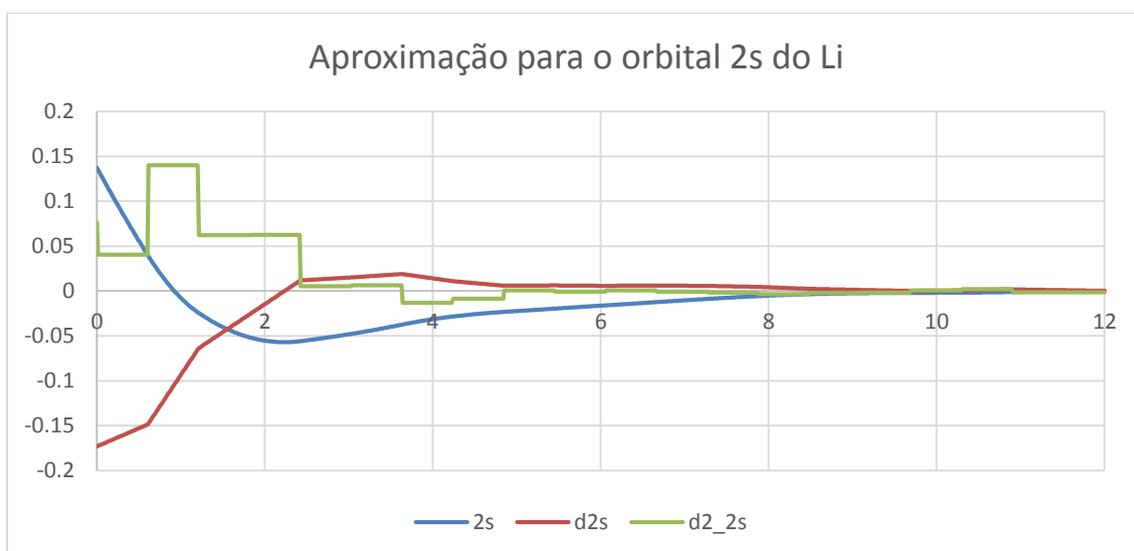


Figura 7-4 a aproximação do orbital 2s do Li, e suas derivadas

Na Figura 7-5 e Figura 7-6 são mostradas as aproximações das densidades eletrônicas do orbital 1s e 2s

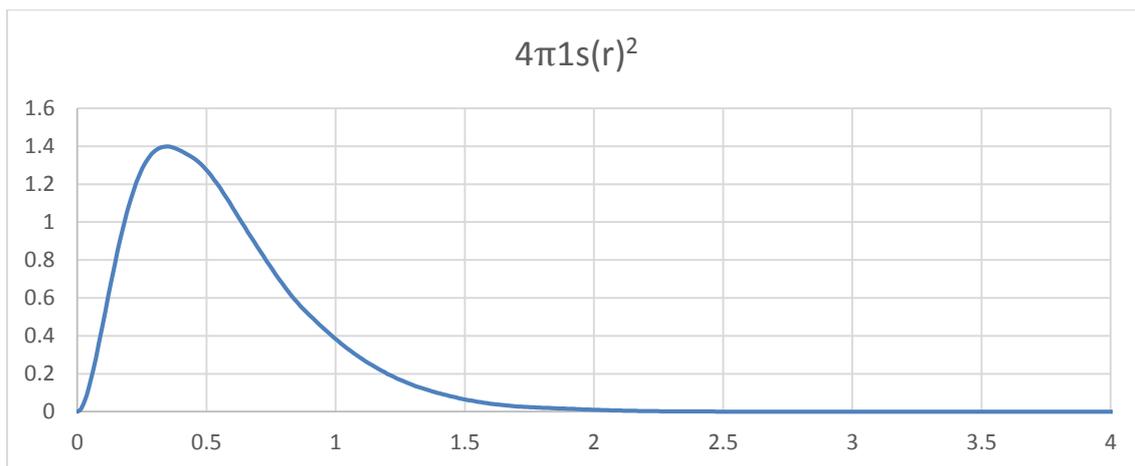


Figura 7-5 Densidade eletrônica representada pelo orbital 1s

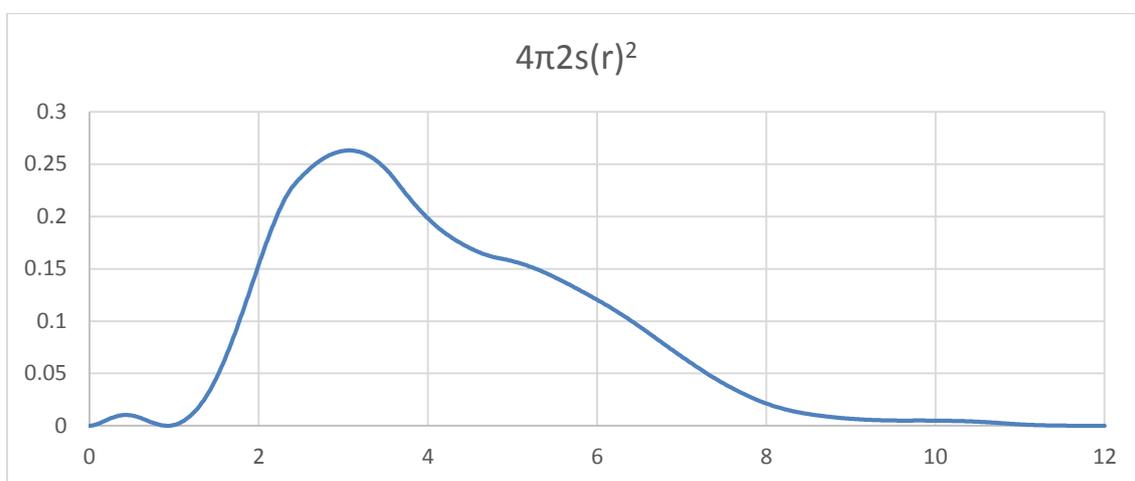


Figura 7-6 Densidade eletrônica representada pelo orbital 2s

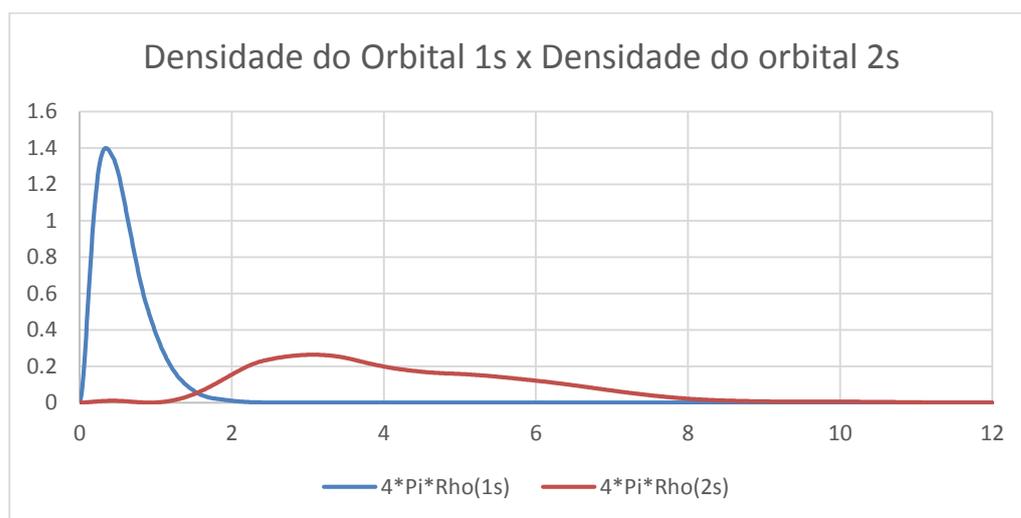


Figura 7-7 Comparação entre os orbitais 1s e 2s do átomo de Li.

7.2.2.**Caso 2b Análise de desempenho dos Algoritmos Co-evolutivos com Inspiração Quântica**

Para o mesmo estudo de caso é realizada uma aplicação dos algoritmos co-evolutivos com inspiração quântica, também desenvolvidos nessa tese, a parametrização do algoritmo está descrita na

Tabela 7-3.

Tabela 7-3 Configurações do AEIQ co-evolutivo utilizado para construir funções de base.

Número de gerações	1000
Número de populações	2
Tamanho população clássica	100
Tamanho da população quântica	2
Número de observações	5
Cruzamento Quântico	0.95
Cruzamento Clássico	0.95
Intervalo de atualização	10 gerações
Variação de largura dos pulsos	0.1

As populações serão responsáveis pelo mapeamento dos fatores multiplicativos em cada argumento, e os parâmetros lineares, respectivamente. As curvas de evolução para os AEIQs e Co-evolutivos estão na Figura 7-8.

Na Figura 7-8 é mostrada uma comparação entre a evolução do AEIQ e sua versão co-evolutiva. Esse experimento é realizado com a finalidade de verificar a se a versão co-evolutiva pode apresentar um desempenho a do AEIQ. E tal como pode ser visto, de fato, os algoritmos co-evolutivo em média apresentam um melhor desempenho, para o caso em estudo.

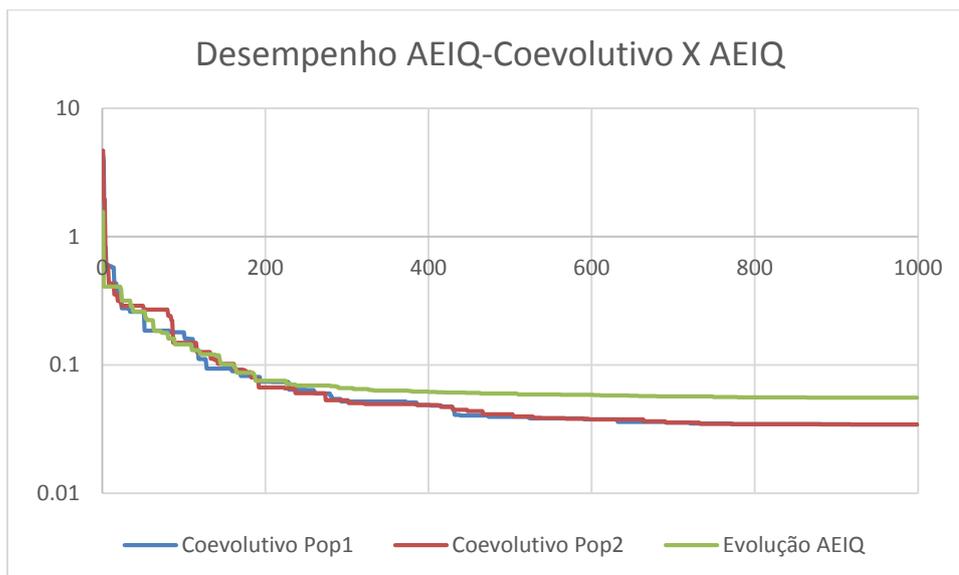


Figura 7-8 Comparação entre o desempenho dos algoritmos AEIQ e sua versão Co-evolutiva.

7.3.

Estudo de caso 3

7.3.1.

Caso 3a - Algoritmos Co-evolutivos com inspiração Quântica e a ortogonalização de Gram-Schmidt

Uma das grandes vantagens decorrentes da utilização de algoritmos evolutivos é a possibilidade de se introduzir codificações específicas ao problema em questão. Além disso, esses algoritmos permitem a incorporação e manipulações específica-se dos cromossomos de cada indivíduo da população. Dessa forma, conforme notado no exemplo anterior a ausência da ortogonalidade entre os spin-orbitais impede que o algoritmo genético alcance valores de energia ainda mais baixos. Logo, naturalmente pode se utilizar um procedimento para tornar as funções representadas pelos cromossomas ortogonais entre si, o procedimento de Gram-Schmidt.

A utilização desse procedimento é feita de forma natural, conforme posto no capítulo 3. Assim, o conjunto de funções codificadas em cada cromossoma é transformado em um conjunto ortonormal seguindo o procedimento de ortogonalização de Gram-Schmidt. Dessa forma, é esperado que o algoritmo aumente a velocidade de convergência, uma vez que não serão necessárias

penalizações. Isso pode ser ilustrado nas curvas de evolução presentes na Figura 7-9.

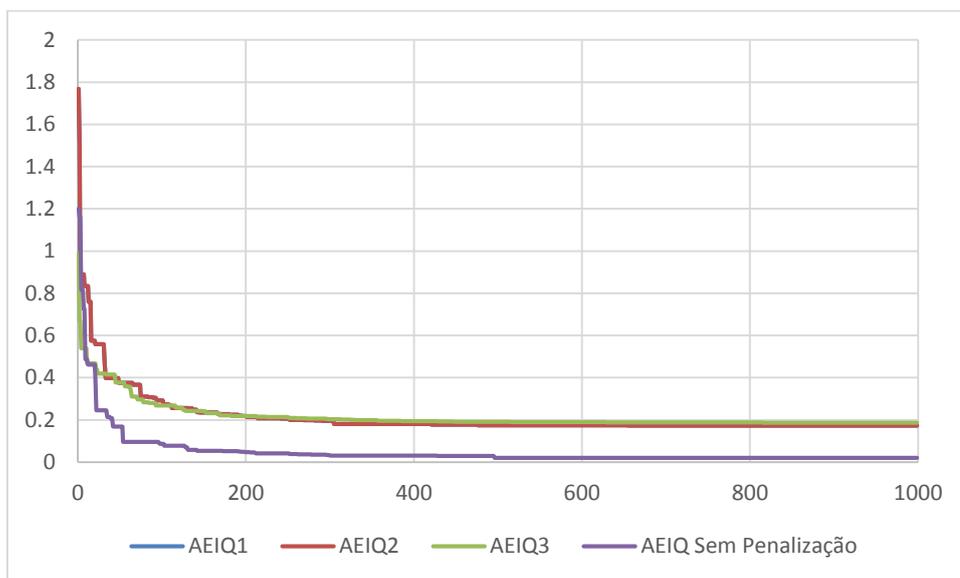


Figura 7-9 Comparação de desempenho entre o método do estudo de caso 1 e estudo de caso 3

Novamente aqui são executados o AEIQ e sua versão co-evolucionária. Foram executados cinco experimentos para o átomo de Li de configuração $1s22s$. No entanto, a energia obtida foi de -7.4244 contra -7.41504 presente no caso 2b.

Na Figura 7-12 Figura 7-13 e Figura 7-14 são feitas comparações entre os orbitais $1s$ e $2s$ e suas respectivas densidades radiais, provenientes do estudo de caso 2b e 3a.

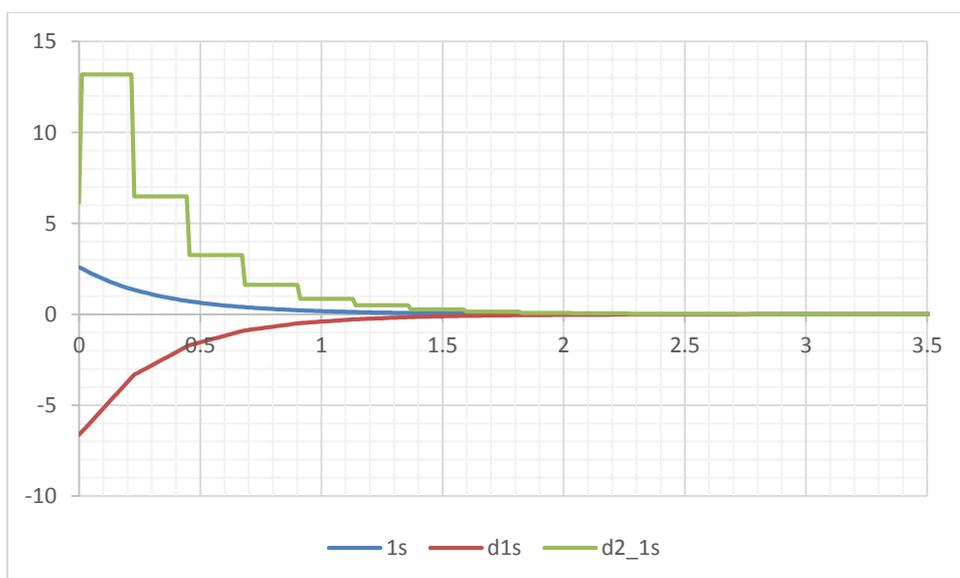


Figura 7-10 Orbitais $1s$, primeira derivada e segunda derivada

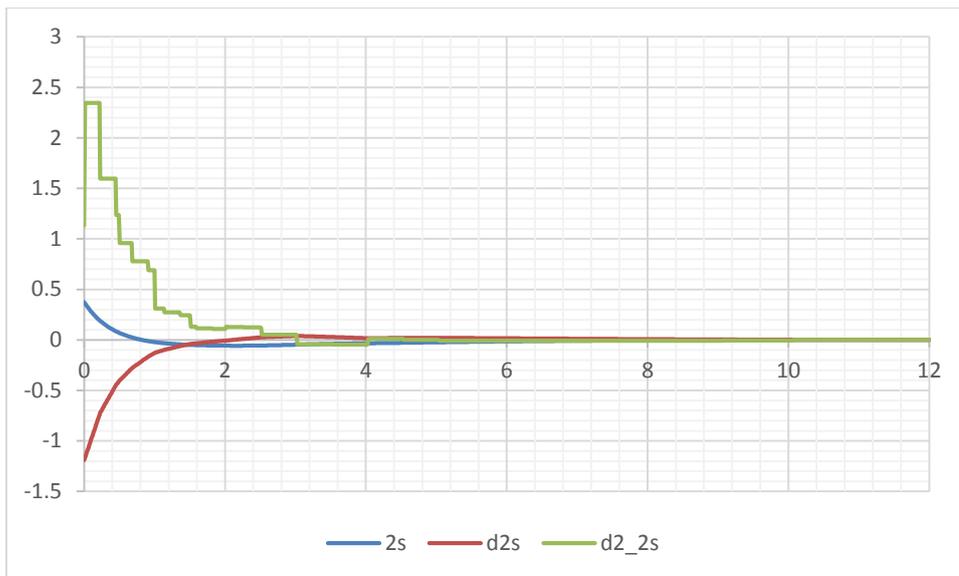


Figura 7-11 Orbitais 2s, primeira derivada e segunda derivada

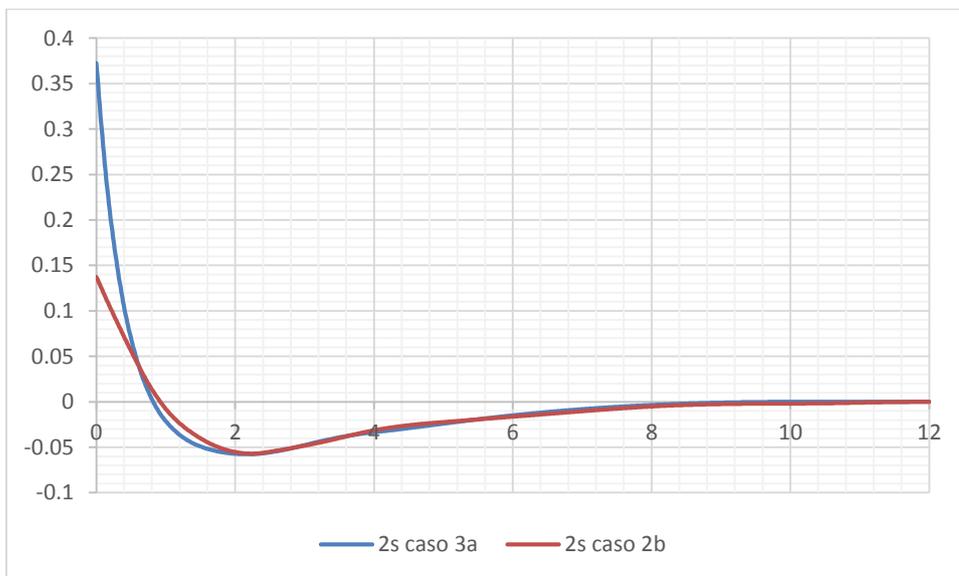


Figura 7-12 Comparação entre o orbitais 2s obtidos no estudo de caso 3a e 2b

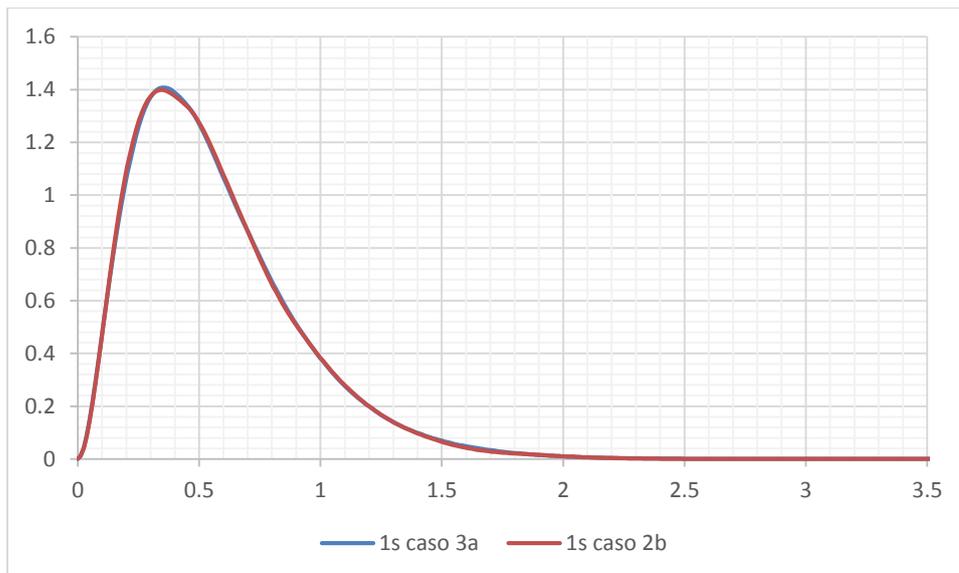


Figura 7-13 Comparação entre as densidades radiais 1s obtidas no estudo de caso 3a e 2b

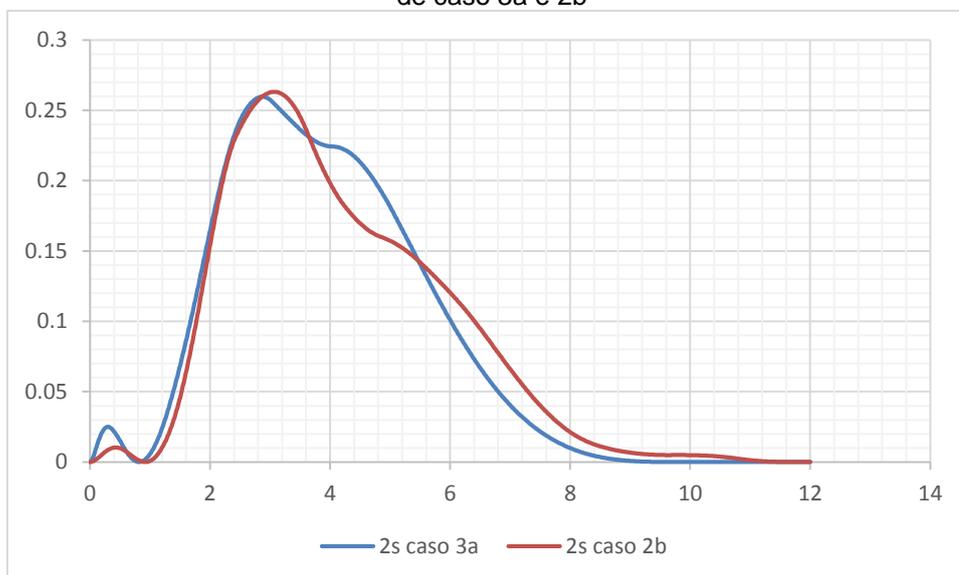


Figura 7-14 Comparação entre as densidades radiais 2s obtidas no estudo de caso 3a e 2b

7.3.2.

Caso 3b Algoritmos Co-evolutivos com inspiração Quântica e a ortogonalização de Gram–Schmidt paramétrico

Conforme o estudo de caso anterior, nota-se que a ortogonalização do conjunto de funções codificadas em cada indivíduo acelera a convergência do algoritmo evolutivo. Existe outra modificação que pode também acelerar a convergência do algoritmo, a diminuição do número de variáveis no problema de

otimização também é um dos fatores que pode ser responsável pelo aumento da velocidade de convergência na evolução de um Algoritmo Evolutivo.

Nesse ponto, cabe notar que para as funções representadas por cada indivíduo serem ortonormais é suficiente que sejam, a princípio, independentes. Uma vez que esse conjunto seja independente, basta a realizar aplicação da ortogonalização de Gram–Schmidt, tal como no estudo de caso 3a.

Para que seja conseguida a independência entre o conjunto de funções é suficiente utilizar os fatores de ponderação diferentes para o argumento de apenas uma função não linear. Como exemplo, seja o conjunto de funções não-lineares, $f: R_+ \rightarrow R$, $g: R_+ \rightarrow R$, $z: R_+ \rightarrow R$, sejam α e β números reais, onde $\alpha \neq \beta$, se $g(x) = f(\alpha x)$ e $z(x) = f(\beta x)$ essas funções serão independentes entre si, e dessa forma basta realizar a aplicação sistemática do procedimento de Gram–Schmidt afim de obter aproximações para todos os spin-orbitais.

O tamanho do cromossoma será restrito a apenas o necessário para realizar a aproximação e uma função spin-orbital, acrescido de tantos parâmetros não lineares quantos forem os spin-orbitais em questão.

Nas Figura 7-15 e Figura 7-17 são mostradas funções radiais de densidade do átomo de Ne para uma base com $V=1$ e $W=2$.

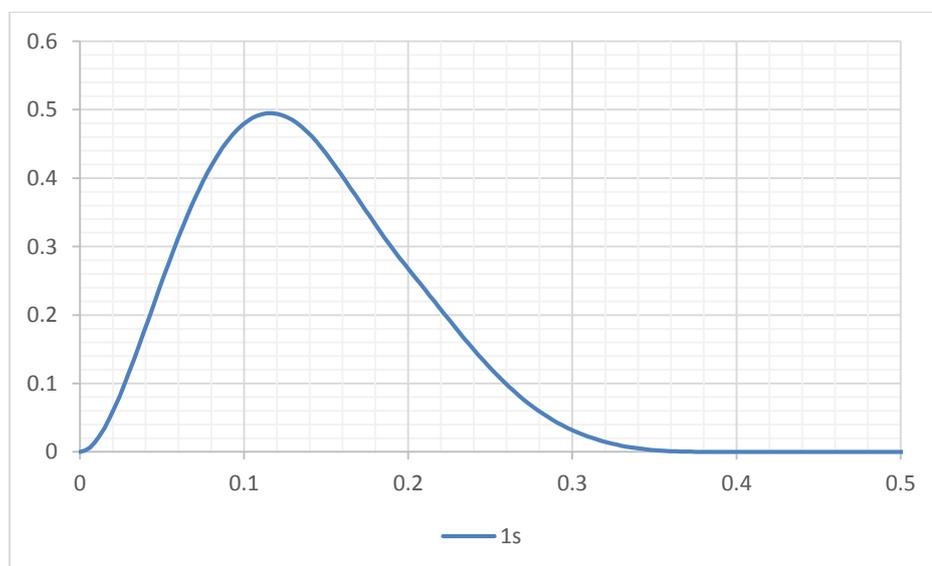


Figura 7-15 Orbital 1s do Ne calculado com a base $V=1$ e $W=2$

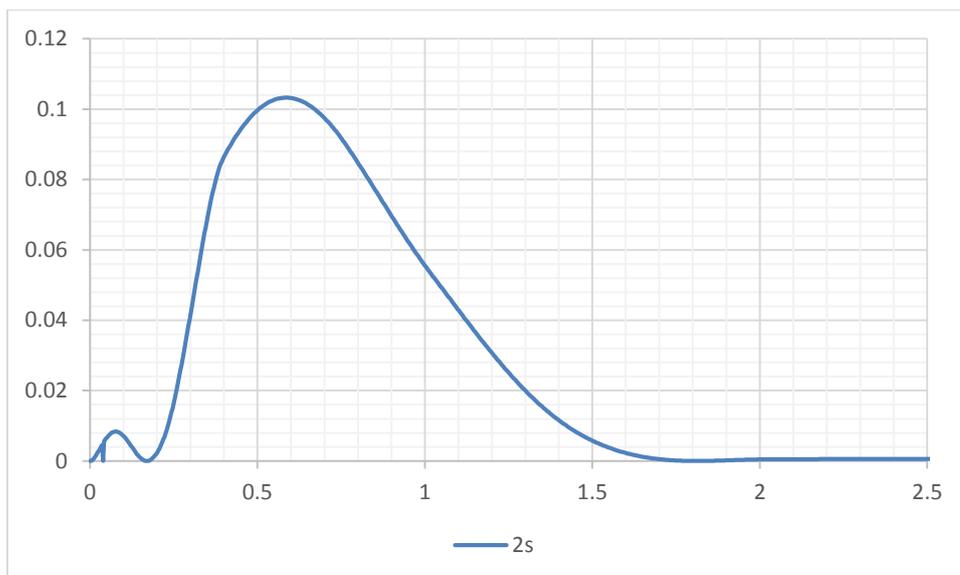


Figura 7-16 Orbital 2s do Ne calculado com a base V=1 e W=2

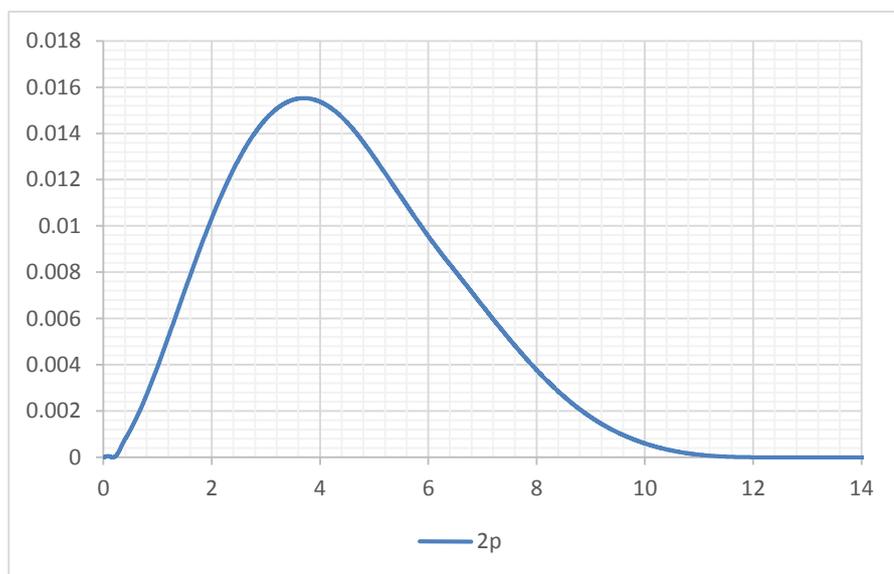


Figura 7-17 Orbital 2p do Ne calculado com a base V=1 e W=2

Tabela 7-4 Base para o Ne com V=1 e W=2

$\alpha \beta \lambda$	V	W
5.018388	-4.334072	0.977419
0.157361	-2.207906	0.129413
0.979467	0.208423	-0.158698
	-0.235522	-0.014997
	-0.310979	0.007068
		0.522006
		0.062804
		0.068612
		0.010280
		0.043001
		0.074936
		0.039495

A energia calculada para o Ne com a base resultante do cromossoma da

Tabela 7-4 é de -128.547099261325, o que representa uma diferença de aproximada 0.001 micro hartree de diferença com relação ao cálculo da energia via Hatree Fock numérico. Esse estudo de caso comprova a viabilidade do emprego das bases sugeridas em cálculos de estrutura eletrônica de alta precisão.

7.4.

Caso de estudo 4

A abordagem desenvolvida no estudo de caso 3b pode ser utilizada para a aproximação dos spin-orbitais de átomos mais complexos. Aqui serão realizadas aproximações para os átomos B, C, N, O, F, Ne. Todo o desenvolvimento aqui é feito considerando funções de onda determinantes de Slater. Todos os spin-orbitais serão o resultado das aproximações sugeridas construídas a partir da base derivada da base de Haar, o procedimento para outras bases propostas também pode ser facilmente repetido.

7.5.

Caso 4a – Utilização da bases para cálculo dos átomos de Li, Be, B, C, N, O , F, Ne

Diferente dos experimentos anteriores são aqui serão calculadas energias para um conjunto de átomos. Para uma base pequena $V=1$ e $W=1$.

São aqui utilizadas as mesmas parametrizações dos estudos de caso anteriores.

Tabela 7-5 Energia de alguns átomos calculada a partir das funções de base com $V=1$, $W=1$

Átomo	Energia
Li	-7.3471
Be	-14.404
B	-24.0496
C	-36.4484
N	-51.97
O	-74.3967

F	-99.4094
Ne	-126.029

Tabela 7-6 Parâmetros das bases para uma aproximação com $V=1$ e $W=1$

					α	β	λ
Li	0.91858	0.044918	0.582653	0.016919	0.867373	0.109609	-0.63424
Be	0.505587	0.027009	0.307528	0.011087	1.207991	0.17226	0.222966
B	0.383296	0.022383	0.224335	0.009738	1.558641	0.262143	0.077645
C	0.421661	0.025014	0.241819	0.01169	1.880527	0.342228	0.115932
N	0.397194	0.025357	0.22037	0.012773	2.224388	0.411893	0.154973
O	0.226315	0.019531	0.097171	4.10E-05	2.980459	0.055457	0.60098
F	0.339374	0.006414	0.219653	-0.00022	2.158174	0.098681	0.428501
NE	0.358521	-0.00619	0.295946	0.001527	1.682076	0.700941	0.192287

CONCLUSÕES

As bases propostas aqui apresentam características que facilitam a implementação computacional. Um bom exemplo são as bases derivadas da base de Haar, que são funções quadráticas por partes. A metodologia desenvolvida para realizar a construção das bases permite que grande parte dos cálculos necessários para a obtenção das energias de sistemas eletrônicos seja feita de forma analítica. Todas as derivadas necessárias, por exemplo, são obtidas analiticamente, além de existir algumas outras propriedades analíticas a serem exploradas, como as integrais de dois elétrons. Como já dito, as bases mais simples propostas aqui, derivadas a partir da base de Haar, são funções quadráticas por partes.

Todo o conjunto de funções de base construído pode ser utilizado juntamente a outros formalismos, como a Teoria do Funcional de Densidade. Isso, permite a simulação de sistemas muito mais complexos. Outro ponto que merece destaque é que os spin orbitais, em qualquer formalismo, são aproximados por bases do L^2 , o que fornece uma garantia da obtenção de resultados assintoticamente corretos, diferentemente das bases de Slater ou Gaussianas tradicionais que não são bases completas para o L^2 .

Com uma base pequena, com poucos parâmetros, mostrou-se que é possível realizar cálculos de alta acurácia. Isso aumenta a velocidade da obtenção de resultados e sua confiabilidade.

A utilização de algoritmos evolutivos e co-evolutivos, conforme visto nos vários estudos de caso, mostrou-se uma ferramenta adequada para a realização de construções das bases. Pois, além de permitirem a construção de bases a partir da otimização de um funcional, essa classe algoritmos é flexível e dinâmica o suficiente para garantir a incorporação das diversas restrições e imposições aos cromossomas, que são os responsáveis pela representação das soluções propostas.

Idealmente, é possível visualizar a utilização de algoritmos evolutivos e co-evolutivos para viabilizar cálculos de estrutura eletrônica que considerem a evolução de bases diferentes cálculos de moléculas específicas, evitando, por exemplo, a combinação linear de orbitais atômicos, LCAO. A função objetivo

seria a mesma, a energia do sistema, e os parâmetros seriam todos os coeficientes, como como no caso de átomos. Isso possibilitaria a construção de soluções mais precisas e confiáveis, porém haveria uma grande quantidade de parâmetros, que é função do número de elétrons do sistema considerado. Essa quantidade de parâmetros considerada, apesar de poder ser considerada um obstáculo, pode ser tratada pela classe de algoritmos sugeridos nessa tese. Pois, essa classe de algoritmos já se mostrou escalável e capaz de tratar problemas com até milhares de parâmetros, (CRUZ, 2007).

Este trabalho propôs a introdução de um método baseado em Algoritmos Evolutivos, para cálculo da estrutura eletrônica de átomos e moléculas, que converge em probabilidade para a resposta correta, capaz de considerar o conjunto $\mathcal{F}^+ = \overline{C^2(0, \infty) \cap L^2(0, \infty) \cap S}$ em $L^2(0, \infty)$ como espaço de soluções viáveis. Esse método consiste na aplicação de um otimizador fundamentado em estratégias estocásticas de otimização aplicada à minimização do *range* numérico do operador Hamiltoniano de sistemas atômicos e moleculares. Essa minimização corresponde ao problema de encontrar o menor autovalor de um operador auto-adjunto. Para que essa tarefa fosse possível, foi necessário aplicar alguns teoremas e definições presentes da Matemática Aplicada. Foram obtidos e compilados alguns resultados que serão fundamentais para a extensão de toda a metodologia desenvolvida, uma vez que, até este momento, foram estudados apenas sistemas de poucos elétrons.

Considera-se como principais resultados dessa tese:

- i. **Uma aplicação inédita do algoritmo co-evolutivo com inspiração quântica;** é importante mencionar que obter o número de base e os seus coeficientes ao mesmo tempo, tornou-se possível através da abordagem evolucionária;
- ii. **Um novo paradigma de utilização de funções wavelets como funções de base para cálculos de estrutura eletrônica;** foi desenvolvido um arcabouço inédito que possibilita utilizar bases completas, muito simples, como a de Haar, por exemplo, em cálculos de estrutura eletrônica, baseado na proposição 4-1;

- iii. **A construção de uma metodologia inédita capaz de construir bases a partir de qualquer base ortogonal de suporte compacto;** foram construídas uma família inédita de bases completas para cálculos de estrutura eletrônica em geral;
- iv. **A construção de uma nova base baseada no segundo polinômio associado de Laguerre modificado;** que pode acelerar a convergência do método proposto para a solução correta. Isso pela inserção de um parâmetro não-linear na própria formação da base, variando o parâmetro da base $q'_n(\cdot, a)$, possibilitando ao algoritmo realizar a variação do sub espaço no qual a solução é procurada;
- v. **Uma base cujos elementos podem ser obtidos por recursividade;** isso possibilita a reutilização dos elementos de base para cálculos mais apurados; dessa forma, é vislumbrada a possibilidade da redução do esforço computacional para a aproximações de ordem superior. Apesar disso, como visto, realizar a expansão de uma função de onda em séries de Fourier pode demandar uma quantidade muito grande de integrais a serem computadas para realizar o cálculo da energia total do sistema;
- vi. **Uma metodologia para garantir que o problema de otimização é um problema restrito;** que é consequência do teorema 6-9 e de seus corolários que estabelecem limites superiores e inferiores para os coeficientes das séries de Fourier;
- vii. **Um teorema que pode contribuir para a resolução de dimensionalidade de um problema de otimização;** esse resultado é o teorema 9.6 dessa tese;
- viii. **As discussões sobre a complexidade computacional do cálculo do range numérico do operador de Schrödinger, e os caminhos sugeridos.**

SUGESTÕES PARA TRABALHOS FUTUROS

A energia total do sistema quântico é uma soma das contribuições das energia cinética, potencial de interação elétron núcleo, e potencial inter-eletrônico. É possível notar nos cálculos das energia cinética e potencial a presença de equações, da forma

$$\left(\int_0^{\infty} f(r_2)^2 r_2^2 dr_2 \right). \quad 9-1$$

É possível reduzir uma parcela significativa de custo computacional, caso a função f seja escrita como uma combinação de bases ortonormais com relação a função peso r^2 . Dessa forma, torna-se possível reestruturar o conjunto de equações, já descrito, por um conjunto de equações ainda mais simples. De fato, se

$$f(r_2) = \sum_{i=1}^{\infty} c_i \varphi_i(r_2) \quad 9-2$$

$$\int_0^{\infty} f(r_2)^2 r_2^2 dr_2 = \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} c_i c_j \int_0^{\infty} \varphi_i(r_2) \varphi_j(r_2) r_2^2 dr_2, \quad 9-3$$

$$\varphi_i(r_2) \perp \varphi_j(r_2) r_2^2, \forall_{i,j}, \quad 9-4$$

Isso implica numa grande redução de cálculo. Isso devido a não ser mais necessário calcular os termos cruzados, isto é, apenas é necessário fazer os cálculos das contribuições dos elementos de bases indexadas sobre o mesmo índice. Tomando o fato das bases serem ortonormais com relação a função r^2 , tem-se

$$\left(\sum_{i=1}^{\infty} \sum_{j=i}^{\infty} c_i c_j \int_0^{\infty} \varphi_i(r_2) \varphi_j(r_2) r_2^2 dr_2 \right) = \sum_{i=1}^{\infty} c_i^2 \quad 9-5$$

Mas, será que,

$$\bigvee_i \varphi_i = (L^2, \|\cdot\|_2), \quad 9-6$$

caso o produto interno considerado seja dado por

$$\int_0^{\infty} \varphi_i(r_2) \varphi_j(r_2) r_2^2 dr_2 ?$$

Um procedimento mais completo, e provavelmente mais sofisticado, seria encontrar um conjunto ortornormal que satisfizesse a seguinte relação,

$$\begin{cases} \int_0^{\infty} \varphi_i(r_2) \varphi_j(r_2) r_2^2 dr_2 = \delta_{i,j} \\ \int_0^{\infty} \varphi_i(r_2) \varphi_j(r_2) dr_2 = \delta_{i,j} \end{cases}, \text{ para todo } i, j \text{ natural.} \quad 9-7$$

No entanto, com tais proposições, surgem novamente outras questões teóricas e práticas a serem respondidas:

- 1) Será que a solução de tal sistema existe?
- 2) Se existe, como seria possível resolvê-lo?
- 3) A solução pode ser obtida por uma modificação do procedimento de Gram-Schmidt?
- 4) Qual seria o custo de tal procedimento?

Com a aplicação dessa metodologia, é possível calcular, previamente, todas as contribuições individuais de cada um dos elementos de base, na energia total do sistema, para qualquer sistema atômico ou molecular. De posse disso, é possível desenvolver uma relação de recursividade entre os termos da sequência obtida após o cálculo. Na figura 6-1 são mostradas duas bases que são ortogonais na semi-reta com relação a função r^2 . A figura 6-1 a) mostra curvas provenientes dos elementos de base dos polinômios de Laguerre do tipo 2, já a figura 6-1 b) ilustra o comportamento assintótico dos cinco primeiros elementos de uma base construída a partir da ortogonalização de *Gram-schmidt* do conjunto definido por $\{e^{nx}\}_{n \in \mathbb{N}}$.

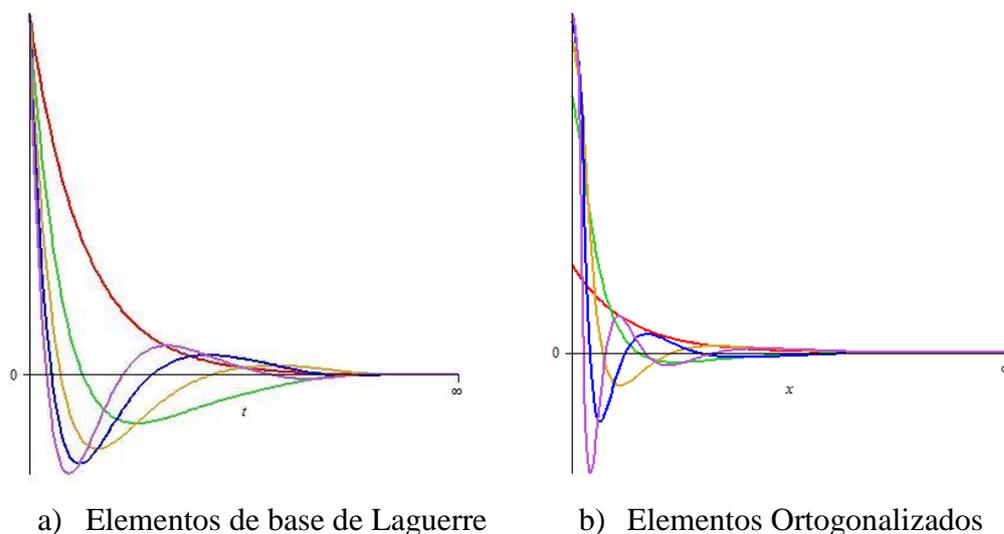


Figura 6.1 – Cinco primeiros elementos da base de Laguerre e da base proveniente do conjunto $\{e^{nx}\}_{n \in \mathbb{N}}$ ortogonalizado por *Gram-Schmidt*.

9.1.1.

Redes Neurais e Neuro-evolução aplicados à solução da equação de Schrödinger

Nos últimos anos, foram realizados grandes avanços na identificação de processos e sistemas não-lineares com a utilização de modelos provenientes da inteligência computacional.

Nessa área destacam-se os modelos obtidos a partir de regras (Lógica Difusa) e, principalmente, os que utilizam funções de ativação (Redes Neurais).

Na visão da inteligência computacional, uma rede neural é composta de camadas de unidades processadoras interconectadas através de pesos e conhecidas por neurônios, na visão da matemática uma rede neural é uma *ridge function*. Um sinal ao passar por um neurônio sofre uma transformação, em geral não-linear, aplicada por uma função de ativação. As Redes Neurais mais utilizadas são as do tipo feed forward. Esse tipo de rede normalmente pode ser representada como mostra a Figura 9-1.

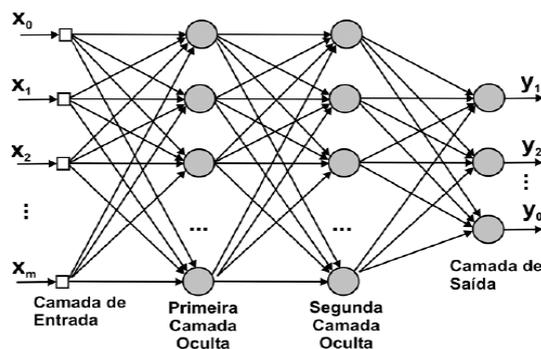


Figura 9-1 Rede neural tipo feed forward.

Tal arquitetura caracteriza-se por um processamento de informações dividido em camadas. Conforme demonstrado (CYBENKO, 1990), é possível aproximar qualquer função contínua com redes compostas de duas camadas: uma intermediária e uma de saída. Nesse caso, as redes feed forward que possuem uma saída e uma camada intermediária, podem ser descritas como:

$$z = g\left(\sum_{i=1}^N w_i \cdot g\left(\sum_{j=1}^M \alpha_{ij} \cdot x_j + b_i\right)\right) + b \quad 9-8$$

onde:

N, M : são os números de neurônios da camada intermediária e de variáveis de entrada, respectivamente;

z : valor de saída;

w_i e $\alpha_{i,j}$: pesos da camada de saída e entrada, respectivamente ($i = 1, \dots, N$) ($j = 1, \dots, M$).

b_i e b : bias da camada intermediária ($i = 1, \dots, N$) e da camada de saída;

g : mapeamento não-linear. Alguns dos mais utilizados são a função sigmoide e a tangente hiperbólica;

$X = [x_1, x_2 \dots x_M]^T$: vetor de variáveis de entrada;

Dessa forma, cabe a utilização de funções da forma sugerida pela equação (4-5), pois segundo (CYBENKO, 1990) e (HAYKIN, 2001), esta forma funcional tem capacidade de aproximar um conjunto bastante grande de funções diferenciáveis e contínuas.

A utilização de Redes Neurais para sintetizar funções de base difere das formas mais comuns de suas aplicações em outros contextos, pois é realizado um procedimento, ao qual se poderia chamar de Neuro-Evolução fracamente supervisionada. Baseados em Algoritmos Evolucionários, são feitos ajustes dos

pesos, isto é, dos parâmetros da rede neural, para que seja efetuada a minimização do sistema em questão. Nesse caso, é possível utilizar uma função descrita por uma rede neural candidata à solução do problema de encontrar o mínimo do range numérico do operador de Schrödinger. A busca por funções do tipo 5.130 é tratada na literatura de inteligência computacional como abordagem Neuro Genética.

Definição 7.1 (*Ridge Function*)

Uma *Ridge Function*, na sua forma mais simples, é uma função multivariada $f: \mathcal{R}^n \rightarrow \mathcal{R}$ como segue

$$f(x_1, x_2 \dots x_n) = g(a_1 x_1 + a_2 x_2 \dots + a_n x_n) = g(a \cdot x),$$

onde $g: \mathcal{R} \rightarrow \mathcal{R}$ e $a \in \mathcal{R}^n \setminus \{0\}$.

Uma rede neural pode ser escrita da seguinte forma

$$R(\mathbf{x}) = \sum_{i=1}^n \lambda_i g(\mathbf{a}^i \cdot \mathbf{x}), \quad 9-9$$

onde $\mathbf{a}^i, \lambda_i \in \mathcal{R}^n \setminus \{0\}$.

Sob certas condições impostas à função g , pode ser mostrado que o conjunto formado por todos os possíveis $R(\mathbf{x})$ é denso em espaços funcionais de grande interesse, como o $(C(0,1), \| \cdot \|_\infty)$.

Uma das sugestões para trabalhos futuros é apresentar um método que possibilite a construção de Redes Neurais aptas a aproximar funções definidas no espaço $\mathcal{F}^+ := \overline{C^2(0,\infty) \cap L^2(0,\infty) \cap S}$ em $L^2(0,\infty)$ na norma $\| \cdot \|_2$.

9.2.

Cálculo Rápido das Integrais Eletrônicas por Inteligência Computacional

As integrais da forma $\int_0^\infty \varphi_i K \varphi_j d\mu$ são funções apenas dos índices i e j , isso para uma base Φ previamente escolhida. Então, pode-se definir uma função $Y: \mathbb{N}^2 \rightarrow \mathcal{R}$ tal que

$$Y(i, j) = \int_0^\infty \varphi_i K \varphi_j d\mu. \quad 9-10$$

Em outras palavras, Y é uma matriz com infinitos termos nas linhas e colunas.

Por inteligência computacional, através de técnicas como Programação Genética (KOZA, 1992) é possível definir heurísticas para especificação dos elementos da matriz Y com base na apresentação de exemplos prévios.

O emprego de tal técnica tornaria o cálculo dos elementos de Y imediato. No entanto, esse cálculo estaria suscetível ao domínio de integridade da heurística de especificação, proveniente da metodologia empregada.

Vale notar que as heurísticas elaboradas por métodos baseados em inteligência computacional não necessariamente podem ou devem ser compreendidas, uma vez que podem resultar em puro código de máquina.

Da mesma forma, poder-se-ia aplicar a mesma idéia já descrita e definir uma função $Y: \mathbb{N}^4 \rightarrow \mathcal{R}$ tal que

$$Y(i, j, k, l) = \iint_0^\infty \varphi_i(\mathbf{x}) \varphi_j(\mathbf{y}) K \varphi_k(\mathbf{x}) \varphi_l(\mathbf{y}) d\mathbf{x} d\mathbf{y}. \quad 9-11$$

9.3.

As Integrais Bi-Eletrônicas

O cômputo da interação elétron-elétron é também conhecido como integral bi-eletrônica, ou mesmo integral de Coulomb, V_{ee} , dado por

$$\langle \Psi(\mathbf{r}_1, \mathbf{r}_2) \left| \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \Psi(\mathbf{r}_1, \mathbf{r}_2) \rangle. \quad 9-12$$

No método proposto, o seu cálculo é dado por

$$\begin{aligned}
\langle \psi(r_1, r_2) | \frac{1}{|r_1 - r_2|} | \psi(r_1, r_2) \rangle \\
= \langle f(r_1)g(r_2) | \frac{1}{|r_1 - r_2|} | f(r_1)g(r_2) \rangle \\
+ 2 \langle f(r_1)g(r_2) | \frac{1}{|r_1 - r_2|} | f(r_2)g(r_1) \rangle \\
+ \langle f(r_2)g(r_1) | \frac{1}{|r_1 - r_2|} | f(r_2)g(r_1) \rangle.
\end{aligned} \tag{9-13}$$

As funções f e g são funções representadas pelas suas respectivas séries de Fourier. Assim, nota-se que é necessário computar as seguintes quantidades:

$$\begin{aligned}
f(r_1)^2 g(r_2)^2 &= \left(\sum_{i=1}^{\infty} c_i \varphi_i(r_1) \right)^2 \left(\sum_{j=1}^{\infty} c'_j \varphi_j(r_2) \right)^2 \\
&= \left(\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} c_i c_j \varphi_i(r_1) \varphi_j(r_1) \right) \left(\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} c'_k c'_l \varphi_k(r_2) \varphi_l(r_2) \right) \\
&= \lim_{n \rightarrow \infty} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n c_i c_j c'_k c'_l \varphi_i(r_1) \varphi_j(r_1) \varphi_k(r_2) \varphi_l(r_2)
\end{aligned} \tag{9-14}$$

$$\begin{aligned}
\langle f(r_1)g(r_2) | \frac{1}{|r_1 - r_2|} | f(r_1)g(r_2) \rangle \\
= \lim_{n \rightarrow \infty} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n c_i c_j c'_k c'_l \iint \varphi_i(r_1) \varphi_j(r_1) \frac{1}{|r_1 - r_2|} \varphi_k(r_2) \varphi_l(r_2) dr_1 dr_2
\end{aligned} \tag{9-15}$$

$$\begin{aligned}
f(r_1)g(r_2)f(r_2)g(r_1) \\
= \left(\sum_{i=1}^{\infty} c_i \varphi_i(r_1) \right) \left(\sum_{j=1}^{\infty} c_j \varphi_j(r_2) \right) \left(\sum_{k=1}^{\infty} c_k \varphi_k(r_2) \right) \left(\sum_{l=1}^{\infty} c_l \varphi_l(r_1) \right) = \\
= \lim_{n \rightarrow \infty} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n c_i c_j c'_k c'_l \iint \varphi_i(r_1) \varphi_j(r_1) \frac{1}{|r_1 - r_2|} \varphi_k(r_2) \varphi_l(r_2) dr_1 dr_2
\end{aligned} \tag{9-16}$$

$$\begin{aligned}
f(r_2)^2 g(r_1)^2 &= \left(\sum_{i=1}^{\infty} c_i \varphi_i(r_2) \right)^2 \left(\sum_{j=1}^{\infty} c_j \varphi_j(r_1) \right)^2 = \\
&= \lim_{n \rightarrow \infty} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n c_i c_j c'_k c'_l \iint \varphi_i(r_1) \varphi_j(r_1) \frac{1}{|r_1 - r_2|} \varphi_k(r_2) \varphi_l(r_2) dr_1 dr_2
\end{aligned} \tag{9-17}$$

Nesse ponto, torna-se fundamental notar que o número de integrais duplas necessárias para calcular o termo V_{ee} , a princípio cresce com n^4 . No entanto, o número de integrais duplas presentes nesse cálculo depende fortemente da base utilizada. Caso seja utilizada uma base polinomial, de Laguerre, por exemplo, o

número de monômios presentes em cada elemento da base é de $n + 1$, sendo n o índice de cada base. Assim, o número de integrais duplas a serem calculadas cresce a uma taxa bem maior do que n^4 . Então, vale a pena investigar a utilização de certos argumentos para diminuir de forma abrupta o custo computacional associado ao cômputo dessas integrais bi-eletrônicas. Pela utilização de algumas propriedades do operador $\frac{1}{|r_1 - r_2|}$, aproveitando o fato de ele ser auto-adjunto, por exemplo.

Uma das vantagens de se utilizar os elementos de base $q'_n(\cdot, a)$, sugerida nessa tese, é o seu caráter recursivo. Isso pode ser utilizado para reduzir o número de operações empregadas no cômputo das integrais, caso sejam utilizados métodos numéricos de quadratura gaussiana (STOER & BURLIRSCH, 1980).

Acredita-se que as relações citadas e desenvolvidas em Balder(1978), podem ser utilizadas para a realização do cálculo mais eficiente dessas integrais e das outras empregadas no cômputo do funcional de energia. As relações são as seguintes:

Lema 6.3

Seja $u: \mathcal{R}^3 \rightarrow \mathcal{R}$, se u é uma função derivável tal que $\int u^2 d\mu = 1$. Então,

$$\int \frac{u^2(y)}{|x-y|} dy \leq 4 \int |\text{grad } u(y)|^2 dy, \forall x \in \mathcal{R}^3 \quad 9-18$$

Lema 6.4

Seja $u: \mathcal{R}^3 \rightarrow \mathcal{R}$, onde u é uma função derivável e seja $q_u: \mathcal{R}^3 \rightarrow \mathcal{R}$ definido por, $q_u(x) = \int \frac{u^2(y)}{|x-y|} dy$. Então, as funções q_u são limitadas por $2\|\text{grad } u\| \|u\|$.

Além desses lemas, também são apresentadas outras relações a respeito do seguinte funcional:

$$E(u) = \int |\text{grad } u(x)|^2 dx - 4 \int \frac{u^2(x)}{|x|} dx + \iint \frac{u^2(x)u^2(y)}{|x-y|} dx dy \quad 9-19$$

Assumindo que as três integrais podem ser representadas respectivamente por $K(u)$, $V(u)$ e $W(u)$, então o funcional $E(u) = K(u) - 4V(u) + W(u)$.

Segundo (Balder, 1978), os seguintes fatos podem ser notados:

- i. As desigualdades abaixo são válidas:

$$0 \leq \int |\text{grad } u(x)|^2 dx = \|\text{grad } u\|^2 \leq \|u\|_1^2 \leq \infty, \quad 9-20$$

$$0 \leq \int \frac{u^2(x)}{|x|} dx = q_u(0) \leq 2\|\text{grad } u\| \|u\| \leq 2\|u\|_1^2 \leq \infty, \quad 9-20$$

$$0 \leq \iint \frac{u^2(x)u^2(y)}{|x-y|} dx dy = \int q_u(x)u^2(x) dx \leq 2\|\text{grad } u\| \|u\|^3 \leq 2\|u\|_1^4 \leq \infty, \quad 9-21$$

ii. Se $u: \mathcal{R}^3 \rightarrow \mathcal{R}$, se u é uma função derivável tal que $\int u^2 d\mu = 1$. Então, $E(u)$ é o valor esperado da energia do átomo de Hélio, caso as funções de onda sejam descritas por $\psi(x,y) = u(x)u(y)$.

iii. Restrito a conjuntos limitados em $L^2(\mathcal{R}^3)$ o funcional de energia $E(u)$ é inferiormente semi-limitado. Pois,

$$\begin{aligned} E(u) &\geq K(u) - 4V(u) \geq \|\text{grad } u\|^2 - 8\|\text{grad } u\| \|u\| \\ &= (\|\text{grad } u\| - 4\|u\|)^2 - 16\|u\|^2 \geq 16\|u\|^2. \end{aligned} \quad 9-22$$

O fato de $E(u)$ ser inferiormente semi-limitado justifica as seguintes definições utilizada no teorema que segue:

Definição 6.5

$$\gamma = \inf\{E(u): u \in L^2(\mathcal{R}^3) \cap C^1(\mathcal{R}^3), \|u\| = 1\}.$$

Definição 6.6

Uma sequência $\{u_n\}_0^n$ de funções em $L^2(\mathcal{R}^3) \cap C^1(\mathcal{R}^3)$ é uma sequência de minimizadores para o funcional E se $\|u_n\| = 1 \forall n \in \mathbb{N}$ e o $\lim_{n \rightarrow \infty} u_n = \gamma$.

Teorema 6.7

Seja $\{u_n\}_0^n$ uma sequência de minimizadores para E . Então, existe uma subsequência $\{u_{n_k}\}_0^k$ de $\{u_n\}_0^n$ e um elemento $u_0 \in L^2(\mathcal{R}^3) \cap C^1(\mathcal{R}^3)$ tal que:

- i. $\|u_0\| = 1, E(u_0) = \lim_{k \rightarrow \infty} E(u_{n_k}) = \gamma;$
- ii. $u_{n_k} \xrightarrow{L^2(\mathcal{R}^3) \cap C^1(\mathcal{R}^3)} u_0;$
- iii. $K(u_0) = \lim_{k \rightarrow \infty} K(u_{n_k}) = -\gamma;$
- iv. $u_0 \in L^2(\mathcal{R}^3) \cap C^2(\mathcal{R}^3) \setminus \{0\}.$

9.4. Teoria do Funcional da Densidade Eletrônica (DFT)

Nesse capítulo são feitas considerações sobre um método já estudado e um dos métodos *ab-initio* mais utilizados da atualidade o método baseado no funcional de densidade eletrônica, DFT (*Density Functional Theory*). Esse método apesar

ligar as teorias semi-clássicas de densidade eletrônica e a mecânica quântica de átomos e moléculas apresenta algumas inconsistências e dificuldades de utilização.

Afim de continuar o trabalho desenvolvido, sugerem-se tópicos:

- i. Estudar quais características a densidade eletrônica $\rho(r)$ deve possuir, uma vez que $\psi \in \mathcal{F}^+ : \overline{L^2(\mathcal{R}_+^n) \cap C^2(\mathcal{R}_+^n) \cap S^+}$;
- ii. Estudar o conjunto originado do item i.
- iii. Estudar a possibilidade de definir novo funcional de densidade eletrônica a partir da definição da equação de Schrödinger para funções de onda Ψ .
- iv. Estudar o problema de minimização em termos de densidade eletrônica funcional obtido no domínio de densidades provenientes de \mathbf{i} .

Determinar a densidade eletrônica, correspondente ao estado fundamental de um sistema multieletrônico, equivale a resolver um problema de minimização de acordo com o teorema 5.1. Esse problema de minimização aparece sob um conjunto de diferentes restrições. De antemão, sem a realização dos estudos propostos, nota-se a inclusão de pelo menos uma restrição a mais:

- A integral sob a densidade eletrônica do sistema tem de corresponder a um número natural, e este é o número de elétrons presentes no sistema estudado;

9.5.

Método das aproximações sucessivas aplicado

Diante dos estudos realizados, também é vislumbrada a possibilidade de realizar uma investigação sobre a aplicação do teorema 3.19 (do ponto fixo de *Banach*) para realizar a construção de uma sequência de potências de um operador da seguinte forma:

$$\lim_{n \rightarrow \infty} T^n(\mathcal{H}\psi) = \varepsilon\psi, \quad 9-23$$

Onde:

T é um operador de \mathcal{F}^+ em \mathcal{F}^+ ;

\mathcal{H} é operador Hamiltoniano;

ψ é uma função em \mathcal{F}^+ ;

ε é uma constante real.

9.6. Possibilidade da redução de dimensionalidade em problemas de otimização

Seja $f: \mathcal{R}^n \rightarrow \mathcal{R}$, $g: \mathcal{R} \rightarrow \mathcal{R}$, Π^n o espaço dos polinômios de grau n de coeficientes complexos, e x^* um ponto de mínimo de f . Como x^* é um ponto em \mathcal{R}^n é possível escrever $x^* = \{x_1^*, x_2^*, \dots, x_n^*\}$ que pode ser representado como $x^* = \bigoplus_{i=1}^n x_i^*$.

Teorema 9.8

Para $n + 1$ pontos suportes (x_i, f_i) , $i = 0, 1, \dots, n$, $x_i \neq x_k$ para $i \neq k$, existe um único polinômio $P \in \Pi^n$ com $P(x_i) = f_i$, $i = 0, 1, \dots, n$.

Prova. Ver (STOER & BURLIRSCH, 1980).

Teorema 9.9

Seja x um ponto em \mathcal{R}^n tal que $x = \bigoplus_{i=1}^n g(i)$. Então, existe um único g em Π^{n-1} tal que $\|x^* - x\| = 0$ e assim,

$$\begin{array}{l} \text{minimizar } f(x) \\ \text{sujeito a } f \in \mathcal{R}^n \end{array} \Rightarrow \begin{array}{l} \text{minimizar } f\left(\bigoplus_{i=1}^n g(i)\right) \\ \text{sujeito a } g \in \Pi^{n-1} \end{array}$$

Prova.

Todo ponto em \mathcal{R}^n pode ser representado por uma seqüência em \mathcal{R} , e toda seqüência em \mathcal{R} , por definição, é uma função dos naturais em \mathcal{R} . Do teorema 6.8 segue que todo par ordenado $(i, f(i))$ pode ser representado exatamente por um único polinômio em $\Pi^n(i, P(i))$ para $i < n$.

O teorema 6.9 garante a existência de um polinômio de grau $n - 1$, onde n é a dimensão do espaço de busca, capaz de mapear o par ordenado (i, x_i) para todo $i < n - 1$ e x_i em \mathcal{R} . Assim, o número de coeficientes desse polinômio é o mesmo da dimensão do espaço em questão. Mas, o teorema não exclui a possibilidade da existência de um polinômio em Π^m onde $m < n - 1$. Esse fato pode ser utilizado para reduzir a dimensionalidade do espaço de busca de um problema de otimização. Pois, a solução do problema

$$\begin{array}{ll} \text{minimizar} & f\left(\bigoplus_{i=1}^n g(i)\right) \\ \text{sujeito a} & \\ g \in \Pi^{n-1} & \end{array} \quad \text{pode ser encontrada em} \quad \begin{array}{ll} \text{minimizar} & f\left(\bigoplus_{i=1}^n g(i)\right) \\ \text{sujeito a} & \\ g' \in \Pi^m & \end{array},$$

onde $m \ll n - 1$, onde m e n são definidos nos naturais, uma vez que $\Pi^m \subseteq \Pi^{n-1}$ para todo $m \leq n - 1$.

REFERÊNCIAS BIBLIOGRÁFICAS

ANDZELM, J.; KLOBUKOWSKI, M.; RADZIO-ANDZELM, E.; SASKI, Y.; TATEWAKI, H. **Gaussian Basis Set for Molecular Calculations**, Elsevier, Amsterdam, 1984.

BUCCI, A. **Emergent Geometric Organization and Informative Dimensions in Coevolutionary Algorithms**, 2007, Tese de doutorado, Brandeis University.

BARTLETT, JR. **Neutron-Proton Interactions**, Phys. Rev. 51, 661, 1937.

BIAYE, M.; KONTÉ, A.; NDAO, A.S.; FAYE, N.A.B; WAGUÉ, A. **The singlet doubly excited (nl)² states of helium-like systems**, *Physica Scripta*, 2005.

BRUENING, E.; BLANCHARD, P. **Mathematical Methods in Physics**, Birkhäuser Boston; 1 edition (October 4, 2002).

CAETANO, C. **Redes Neurais Artificiais e Algoritmo Genético no Estudo de Sistemas Quânticos**, 2003, Dissertação (Mestrado em Física) - Instituto Tecnológico de Aeronáutica, ITA, Brasil.

COUTINHO, N.H.M.K. **Métodos de Química Teórica e Modelagem Molecular**, Editora Livraria da Física, 2007.

CRUZ, A.V.A.; BARBOSA, C.R.H.; PACHECO, M.A.C.; VELLASCO, M.B. R. **Quantum-inspired evolutionary algorithms and its application to numerical optimization problems**. In: ICONIP, Berlin, Germany, Springer-Verlag, 2005.

BURRUS, C.S.; GOPINATH, R.A.; H. GUO, H. **Introduction to Wavelets and Wavelet Transforms: A Primer**, Prentice-Hall, ISBN 0-13-489600-9, 1988.

LEVAN, N.; KUBRUSLY, C.S. **Multiresolution approximation scale and time-shift subspaces**, *Multidimens. Syst. Signal Process*, 17, 2006.

CYBENKO, G. **Approximation by superpositions of a sigmoidal function**. *Mathematics of Control, Signals, and Systems, Godalming*, v. 2, n. 4, p. 303-314, 1989.

DIRAC, P.A.M. **Quantum mechanics of many-electron systems**. *Proc. R. Soc. A.*, 123, 714, 1929.

D. FELLER, E.R. DAVIDSON. **Basis Sets for Ab Initio Molecular Orbital Calculations and Intermolecular Interactions**. *Reviews in Computational Chemistry*, New York, pp. 1-43, 1990.

E. HERNANDEZ, E. and G. WEISS. **A First Course on Wavelets**, CRC Press, Boca Raton, 1996.

- FRIESECKE, G. **The multiconfiguration equations for atoms and molecules: Charge quantization and existence of solutions.** *Arch. Rational Mech. Anal.*, 169, 35, 2003.
- FROESE-FISCHER, C. **The Hartree-Fock method for atoms**, Wiley: New York, 1977.
- GOMES, A.S.P. **Avaliação de métodos de integração numérica na discretização de transformadas Integrais**, Dissertação de mestrado, UNICAMP, 2001.
- HAN, K. and KIM, J.-H., **Genetic quantum algorithm and its application to combinatorial optimization problems**, Proceedings of the IEEE Conference on Evolutionary Computation, Piscataway: IEEE Press, 1354-1360, 2000.
- HARTREE, D.R. **The Calculation of Atomic Structures**, John Wiley & Sons, New York, 1957.
- HAYKIN, S. **Neural Networks – A Comprehensive Foundation**. Prentice Hall, 2 edition, 2001.
- HUZINAGA, H.; KLOBUKOWSKI, M.; TATEWAKI, H. **Comment on Dunning's correlation-consistent basis sets**, *Can. J. Chem.*, 63: 1812, 1985.
- HYLLERAAS, Z. **Phys. 54, Neue Berechnung der Energie des Heeliums im Grundzustande, sowie tiefsten Terms von Ortho-Helium**, 1929.
- LÖWDIN, P.O. **Studies of Atomic Self-Consistent Fields. I. Calculation of Slater Functions**, *Phys. Rev.*, 90, 120, 1953.
- I. DAUBECHIES. **Ten Lectures on Wavelets**. *CBMS-NSF Regional Conference Series in Applied Mathematics*, SIAM, Philadelphia, PA, vol. 61, 1992.
- I. DAUBECHIES. **Orthonormal Bases of Compactly Supported Wavelets**, *Comm. Pure Appl. Math.*, vol. 41, 1988
- JENSEN, F. **Introduction to Computational Chemistry**, Wiley: New York, 1999.
- KINOSHITA. **Ground State of the Helium Atom**, *Phys. Rev.*, 105, 1490, 1957.
- KOZA, J.R. **Genetic Programming: On the Programming of Computers by Means of Natural Selection**, *MIT Press*, Cambridge, MA, USA, 1992.
- FRANKOWSKI, K. and PEKERIS, C.L. **Logarithmic Terms in the Wave Functions of the Ground State of Two-Electron Atoms**. *Phys. Rev.*, 150, 366, 1966.
- KUBRUSLY, C.S. **Elements of Operator Theory**, Birkhäuser, 2001.
- LEACH, A. **Molecular Modelling: Principles and Applications**. Prentice Hall, 2001.
- MALLAT, S.G. **Multiresolution approximations and wavelet orthonormal bases**. *Trans. Amer. Math. Soc.*, 315, 69-87, 1989.

MICHALEWICZ, Z. **Genetic algorithms + data structures = evolution programs, 2nd, extended ed.**. Springer-Verlag New York, Inc., New York, NY, USA, 1994.

MOORE, WALTER J. **Physical Chemistry**, Prentice-Hall, Inc., New Jersey, EUA, 1972.

NAKASHIMA, H1, NAKATSUJI, H. **Solving the Schrodinger equation for helium atom and its isoelectronic ions with the free iterative complement interaction (ICI) method.** *J. Chem. Phys.* 2007

NARAYANAN, MOORE. **Quantum-inspired genetic algorithms.** In: Proceedings of IEEE International Conference on Evolutionary Computation, 1996, pp. 61-66.

PAGAMISSE, A. **Discriminação de Texturas pela Transformada Wavelet.** Tese, Instituto de Matemática, Estatística e Computação, UNICAMP, Campinas, 2004.

ROOTHAAN, C.C.J. **New Developments in Molecular Orbital Theory.** *Rev. Mod. Phys.* 1951, 23: 69.

SANTOS, M.G. **Desenvolvimento de funções de base GTO: Aplicação do método da coordenada Geradora através da técnica de discretização Integral Otimizada.** Dissertação de Mestrado, UNICAMP, 1992.

SHAVITTI. **Methods in Computational Physics**, Eds B. Alder, S., New York: Academic Press, vol. 11, 1963.

SZABO, A.; OSTLUND, N. **Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory.** *Courier Dover Publications*, 1996.

TESCHL. **Mathematical Methods in Quantum Mechanics with Applications to Schrödinger Operators**, *American Mathematical Society, Graduate Studies in Mathematics*, vol. 99, 2009.

VIANNA, J.D.M.; FAZZIO, A.; CANUTO, S. **Teoria Quântica de Moléculas e Sólidos**, Editora Livraria da Física, São Paulo, 2004.

SCHWARTZ. **Experiment and Theory in Computations of the He Atom Ground State**, *Int. J. Mod. Phys. E* 15, 877, 2006.

TOSHIO FUKUDA, **Proceedings of IEEE International Conference on Evolutionary Computation**, *Piscataway: IEEE Press*, 61-66, 1996.

Anexo 1

Limites Hartree Fock

Z	Átomo	Multiplicidade	Energia em hartrees	Z	Átomo	Multiplicidade	Energia em hartrees
3	Li	2	-7.432726927	24	Cr	7	-1043.356376000
4	Be	1	-14.573023160	25	Mn	6	-1149.866251000
5	B	2	-24.529060720	26	Fe	5	-1262.443665000
6	C	3	-37.688618950	27	Co	4	-1381.414553000
7	N	4	-54.400934190	28	Ni	3	-1506.870908000
8	O	3	-74.809398450	29	Cu	2	-1638.963742000
9	F	2	-99.409349330	30	Zn	1	-1777.848116000
10	Ne	1	-128.547098000	31	Ga	2	-1923.261009000
11	Na	2	-161.858911600	32	Ge	3	-2075.359733000
12	Mg	1	-199.614636300	33	As	4	-2234.238654000
13	Al	2	-241.876707200	34	Se	3	-2399.867611000
14	Si	3	-288.854362400	35	Br	2	-2572.441332000
15	P	4	-340.718780800	36	Kr	1	-2752.054977000
16	S	3	-397.504895800	37	Rb	2	-2938.357453000
17	Cl	2	-459.482072100	38	Sr	1	-3131.545686000
18	Ar	1	-526.817512600	39	Y	2	-3331.684169000
19	K	2	-599.164786500	40	Zr	3	-3538.995064000
20	Ca	1	-676.758185700	41	Nb	6	-3753.597727000
21	Sc	2	-759.735717800	42	Mo	7	-3975.549499000
22	Ti	3	-848.405996700	43	Tc	6	-4204.788736000
23	V	4	-942.884337400	44	Ru	5	-4441.539487000