

### 3

## Construções de funções de base no modelo de Hartree-Fock

### 3.1.

#### Especificação e otimização de funções de base

Na construção de uma função de base os expoentes das primitivas gaussianas são variados até que a energia do átomo representado seja mínima. Em alguns casos simples, os expoentes são otimizados individualmente. Em outros, os expoentes estão relacionados uns aos outros por alguma equação e os parâmetros destas equações são otimizados. As primitivas obtidas descrevem átomos isolados e não podem descrever com precisão as deformações dos orbitais atômicos devido à presença de outros átomos em uma molécula. Para cálculos moleculares, as primitivas gaussianas devem ser contraídas, isto é, uma combinação linear destas primitivas é usada como uma função de base. Tais funções de base terão coeficientes e expoentes fixos. As contrações são normalmente chamadas de Orbitais Contraídos do Tipo Gaussianas. Neste trabalho serão obtidas primitivas gaussianas, logo as funções de base contraídas não serão detalhadas. Mais informações sobre a contração de funções de base podem ser obtidas em (SZABO, 1996).

O tempo necessário para executar um cálculo de orbital molecular aumenta proporcionalmente à razão de  $N^4$ , onde  $N$  é o número de funções de base. Portanto, aumentar o número de funções de base aumenta significativamente o tempo de cálculo.

### 3.2.

#### Introdução à otimização

É surpreendente o desenvolvimento, nos últimos 20 anos da matemática e da engenharia aplicadas diretamente a problemas de tomada de decisão. Essa tendência foi motivada principalmente pelos benefícios econômicos resultantes de

decisões apropriadas. No contexto de funções de base, a principal motivação é a redução do custo computacional associada à precisão dos resultados. Esse problema pode ser visto como um problema de tomada decisão.

O conceito de melhor ou decisão ótima emerge como uma aproximação fundamental para a formulação de problemas de decisão. Nessa abordagem uma variável real que quantifica o valor da decisão, é isolada e otimizada (maximizada ou minimizada dependendo da situação) por seleções apropriadas entre as alternativas viáveis.

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 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.

### 3.2.1.

#### Fundamentos de programação não-linear

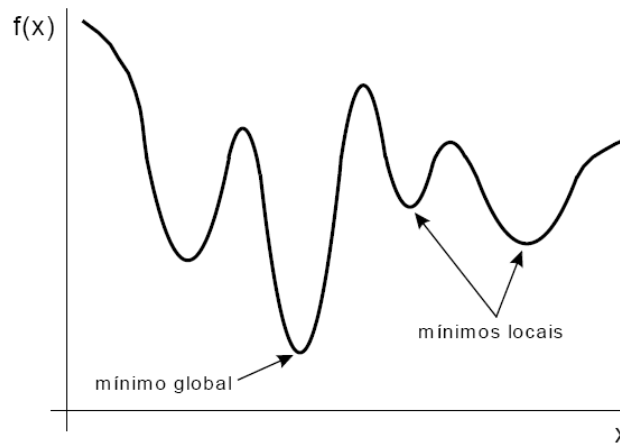
Seja uma função  $f: S \rightarrow \mathbb{R}$ , seja  $x^*, y \in S$ ,  $x^*$  será um ponto de mínimo de  $f$  se existir uma região com raio  $\varepsilon$  em torno de  $x^*$  tal que

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

(3-1)

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 3-1.

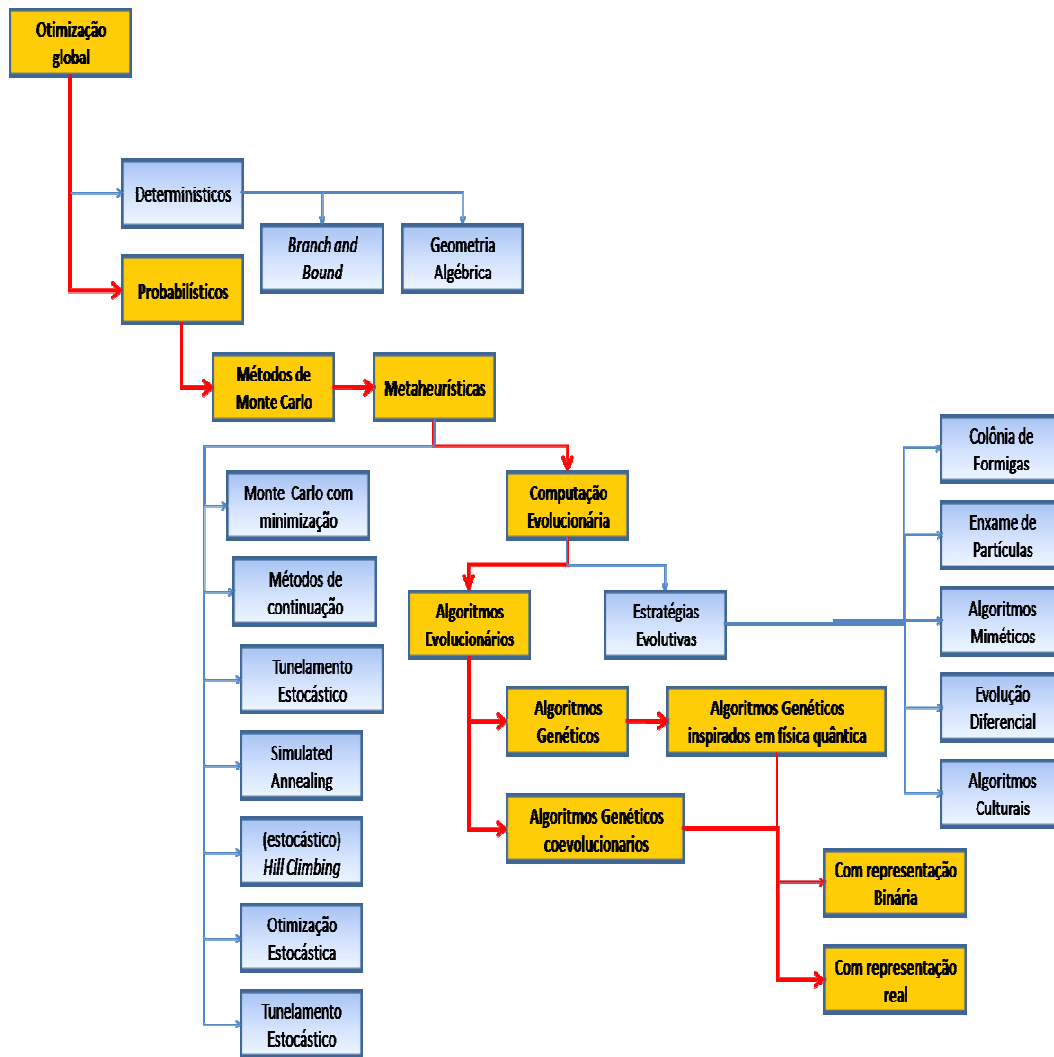


**Figura 3-1** Ilustra uma função que possui mínimos locais e globais.

### 3.2.2.

#### 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 3-2 é mostrada uma taxonomia dos métodos de otimização globais determinísticos e estocásticos.



**Figura 3-2** Taxonomia dos métodos de otimização global, determinísticos e estocásticos.

Nesse trabalho, são particularmente importantes os Algoritmos Evolucionários. Tais algoritmos têm sido extensivamente utilizados na literatura para resolver problemas classificados como complexos. 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.

### 3.2.3.

#### **Otimização multiobjetivo e distância ao alvo**

Quase todos os problemas do mundo real envolvem otimização simultânea sobre objetivos múltiplos e conflitantes (FONSECA & FLEMING, 1995; ZITZLER & THIELE, 1998; COELLO, 1999; *van* VELDHUIZEN & LAMONT, 2000; COELLO, 2001; ZITZLER, 2002). Otimização multiobjetivo, é sem dúvida, um importante tópico de pesquisa para cientistas e engenheiros, não somente por causa da natureza multiobjetiva da maioria dos problemas do mundo real, mas, também, porque ainda existem muitas questões em aberto nesta área (COELLO 1999; ZITZLER, 1999).

Infelizmente, no entanto, é possível verificar que há algumas dificuldades em se aplicar técnicas de otimização a problemas deste tipo, tanto técnicas convencionais, como outras. De fato, a principal dificuldade reside no problema de que, de forma geral, tais técnicas são originalmente projetadas para problemas de um único objetivo.

De acordo com (BLEULER *et al*, 2003), problemas de otimização complexos podem ser achados em muitas áreas de aplicação. Um aspecto que contribui para a complexidade desses problemas compreende as características do

espaço de busca; algoritmos exatos não são aplicáveis. Múltiplos objetivos introduzem outros tipos de dificuldades que métodos de otimização clássicos não foram projetados para resolver. Assim, técnicas alternativas, como algoritmos de busca estocástica têm sido aplicadas.

Quando são aplicados Algoritmos Evolucionários a problemas cuja função objetivo é dada pela agregação de funções objetivo, esses algoritmos ganham um nome especial, Algoritmos Evolucionários Multiobjetivos.

Na prática, é bastante frequente encontrar problemas do mundo real em que uma boa solução deve necessariamente atender satisfatoriamente a todos os objetivos em questão. Nestas situações, não se considera aceitável uma solução que apresente uma avaliação espetacular para um objetivo e que seja ruim para outra. Uma forma de se atingir este objetivo consiste em se avaliar uma determinada solução com  $f$  através da distância entre o vetor formado pelas avaliações  $f_i$  e um vetor-alvo ideal formado pelas avaliações ideais para cada objetivo em separado. Formalmente, esta avaliação pode ser descrita por:

$$f = (\sum_{i=0}^n |\text{ideal} - f_i|^p)^{1/p} \quad (3-2)$$

Desta forma, para  $p=1$ , tem-se a chamada distância Manhattan ou distância metropolitana, que consiste de uma agregação linear de objetivos combinada com uma solução-alvo. Utilizando-se  $p=2$ , entretanto, obtém-se a distância mais comumente utilizada, a distância euclideana. De fato, é possível perceber claramente que a avaliação de soluções torna-se neste caso não-linear, de modo que melhorias numa avaliação  $f_i$  para um objetivo não mais contrabalanceiam, de forma equivalente, pioras na avaliação  $f_i$  para outro objetivo. Efetivamente, a forma quadrática faz com que uma solução não balanceada seja mais penalizada por apresentar um valor  $f_i$  distante do alvo **ideal**, do que beneficiada por ter outro valor  $f_i$  distante de seu alvo **ideal**. Desta forma, pode-se perceber que se exerce agora uma “necessidade de balanceamento”, de modo que se torna mais difícil que uma solução não balanceada seja considerada superior em relação à outra solução mais equilibrada.

Quanto maior o valor de  $p$  utilizado, maior será a penalidade exercida. Em outras palavras, quanto maior o valor de  $p$ , maiores serão as penalizações dadas as soluções que apresentem desempenhos medíocres para algum objetivo. Em especial, para o caso extremo de  $p \rightarrow \infty$ , obtém-se a técnica denominada Minimax

ou MinMax, na qual a avaliação de uma solução corresponde à distância máxima de algum dos  $n$  objetivos em relação ao seu alvo, isto é:

$$\mathbf{f} = \sum_{i=0}^n \mathbf{max}(|\mathbf{ideal}_i - \mathbf{f}_i|) \quad (3-3)$$

Desta forma, a solução ótima corresponderá, no caso de dois objetivos, ao ponto de interseção entre as curvas de avaliação de cada objetivo, não sendo tolerado nenhum desvio deste ponto de equilíbrio.

### 3.2.4.

#### **Fundamentos de programação evolucionária**

##### **Algoritmos genéticos**

Algoritmos Genéticos (AGs) pertencem à classe dos Algoritmos Evolutivos (AEs), 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, pode ser o número de gerações ou até mesmo o tempo de processamento.

##### **Operadores e Parâmetros Genéticos**

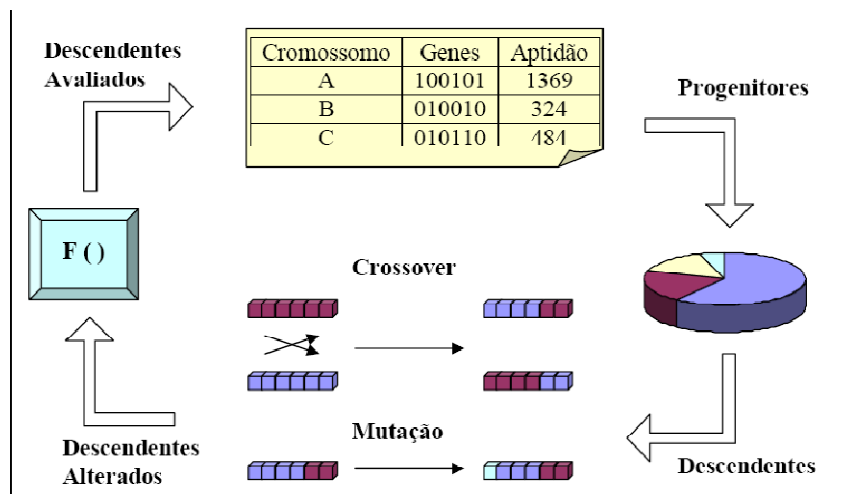
Os operadores genéticos têm a função de, por meio de um processo iterativo, transformar a população inicial em uma população que contenha um

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 3-3 ilustra o esquema de evolução dos AGs Clássicos.



**Figura 3-3** Esquema de evoluçao dos AGs Clássicos.

Apesar de sua grande facilidade de utilizaçao, a análise dos AGs necessita de ferramentas sofisticadas. Pode-se dizer que em Computaçao Evolucionária a convergência de AGs é um dos problemas teóricos mais desafiadores. Muitos pesquisadores exploraram o problema de diversas perspectivas.

Uma das possíveis aproximações para explicar a propriedade de convergência dos AGs é baseada no teorema dos pontos fixos de Banach (KUBRUSLY, 2000). Ele fornece uma visao intuitiva sobre a convergência dos



AGs sem o modelo elitista. O único requisito é que as soluções existentes sejam melhoradas a cada geração (não necessariamente o melhor indivíduo). O Teorema do ponto fixo de Banach promove contrações de mapeamentos em espaços métricos (KUBRUSLY, 2000). Pontos fixos são geralmente aceitos como poderosas ferramentas para definir semântica na computação. AGs podem ser vistos como transformações entre as populações, supondo que seja possível encontrar espaços métricos, nos quais as transformações sejam contrativas. Nesse caso, visualizam-se AGs como pontos fixos sobre o efeito de transformações. Desde que a transformação contenha um único ponto fixo, é conseguida a convergência de AGs como um simples corolário.

### 3.2.5.

#### **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, 2004), (CRUZ, 2007), (NARAYANAN & MOORE, 1996), (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 a grande maioria dos Algoritmos Evolucionários existentes, em conjunto de casos publicados.

### 3.3. 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 3-4 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 3-4** 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):

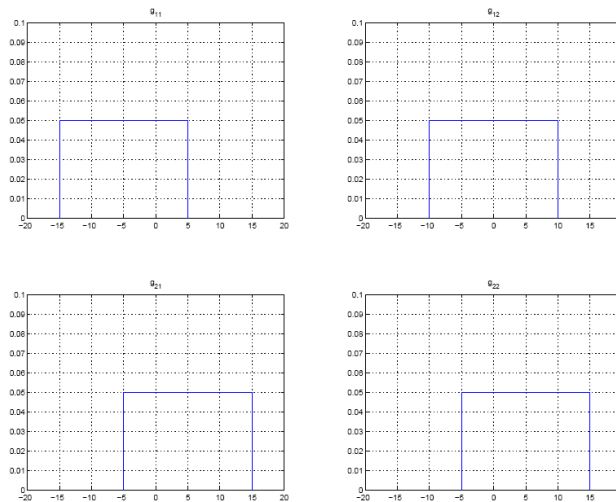
- *Menor tempo para convergência;*
- *O conhecimento sobre o problema que está sendo otimizado é armazenado diretamente nos cromossomos quânticos;*
- *Capacidade de convergir de forma eficiente, mesmo com populações com poucos indivíduos.*

### 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 (3-4) representa um indivíduo de uma população quântica e a Figura 3-5 mostra graficamente a representação de genes quânticos de um indivíduo.

$$Q_i = [p_{ij}, p_{ij+1}, p_{ij+2} \dots] \tag{3-4}$$

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 de funções densidade de probabilidade são usadas pelo AEIQ para gerar os valores para os genes dos indivíduos clássicos.

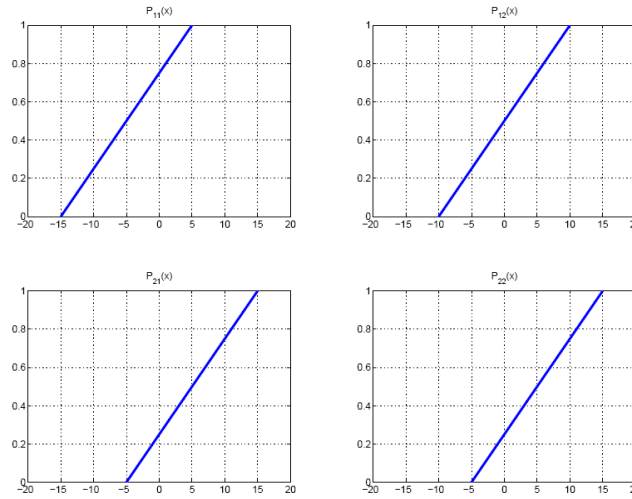


**Figura 3-5** Representação de um indivíduo composto por funções de densidade de probabilidade uniformes.

### 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 um 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 3-6 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 3-5 para o caso implementado em (CRUZ, 2007).



**Figura 3-6** 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 (3-5) é uma das funções mais simples de ser implementadas computacionalmente.

$$p_{ij}(x) = \begin{cases} \frac{1}{U_{ij}-L_{ij}} & \text{se } L_{ij} \leq x \leq U_{ij} \\ 0 & \text{caso contrário} \end{cases} \quad (3-5)$$

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,

$$x_{mj} = r_{mj} * \sigma_{iz} + \left( \mu_{ij} - \frac{\sigma_{ij}}{2} \right) \quad (3-6)$$

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  $\mu_{i1}$  e  $\mu_{i2}$  indicam o centro e  $\sigma_{i1}$  e  $\sigma_{i2}$  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 3-4. 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, cria-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 aderida 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 (3-7)

$$Q(0) = \begin{bmatrix} q_1 = [(\mu = 0, \sigma = 200); (\mu = 0, \sigma = 200)] \\ q_2 = [(\mu = 0, \sigma = 200); (\mu = 0, \sigma = 200)] \\ q_3 = [(\mu = 0, \sigma = 200); (\mu = 0, \sigma = 200)] \\ q_4 = [(\mu = 0, \sigma = 200); (\mu = 0, \sigma = 200)] \\ q_5 = [(\mu = 0, \sigma = 200); (\mu = 0, \sigma = 200)] \end{bmatrix} \quad (3-7)$$

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 3-4) 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 3-4). 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 (3-8).

$$\sigma_{ij} = \left\{ \begin{array}{ll} \sigma_{ij} * \delta & \text{se } \varphi < \frac{1}{5} \\ \frac{\sigma_{ij}}{\delta} & \text{se } \varphi > \frac{1}{5} \\ \sigma_{ij} \text{ caso contrário} & \varphi = \frac{1}{5} \end{array} \right\} \quad (3-8)$$

onde  $\sigma_{ij}$  é a largura do  $j$ -ésimo gene do  $i$ -ésimo indivíduo quântico em  $Q(t)$ ,  $\delta$  é um valor arbitrário, em geral no intervalo  $]0, 1[$  e  $\varphi$  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.

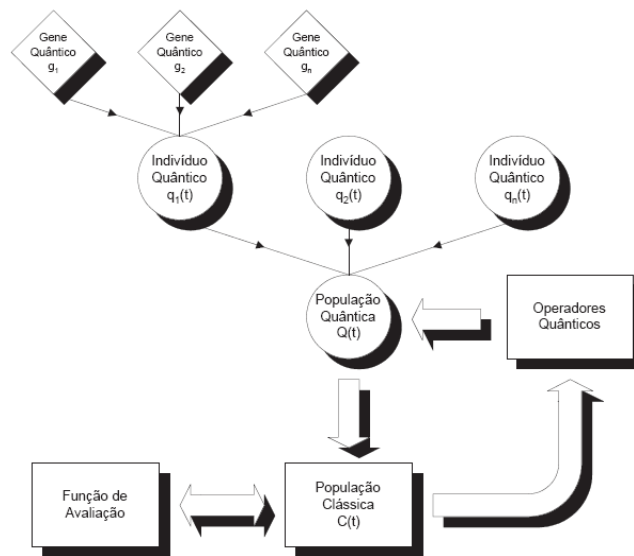
Já no passo dois da atualização dos genes quânticos, o processo pode ser definido em termos de quais indivíduos da população clássica serão usados para atualizar os pulsos da população quântica. Podem ser escolhidos os indivíduos mais aptos, os indivíduos menos aptos, um conjunto aleatório de indivíduos 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  $\lambda$  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 3-4. 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 3-7. Essa figura ilustra graficamente em ordem cronológica, os acontecimentos que ocorrem durante a evolução do algoritmo.



**Figura 3-7** Diagrama completo do Sistema Evolutivo com Inspiração Quântica

O diagrama da Figura 3-7 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.

### 3.4.

#### Um Estudo Bibliográfico sobre a Especificação e Otimização de Funções de base

O desenvolvimento de funções de base é requisito essencial para simulações em química computacional. Como já mencionado anteriormente, qualquer função pode ser uma função de base, mas nem todas possuem características que possibilitam agilizar o cálculo de propriedades atômicas. As funções do tipo Slater e Gaussiana são as que apresentam maior destaque nessa área.

A utilização de funções do tipo Slater como conjunto de base foi discutida em (CLEMENTI *et al*, 1962) e a partir daí surgiram vários trabalhos na mesma linha. Mas o estado da arte da construção de funções de base foi modificado em (HUZINAGA, 1965), que introduziu o uso de funções do tipo gaussianas para o mesmo fim. O uso dessas funções reduziu o esforço computacional para efetuar cálculos de estruturas eletrônicas. Esse trabalho emprega algoritmos de otimização local que minimizam a energia do sistema descrito em função da parametrização das funções de base utilizadas. São diversos os trabalhos que empregam técnicas semelhantes e possuem o mesmo objetivo (PARTRIDGE, 1989), (CHATTOPADHYAYFS, 1981).

Os trabalhos citados anteriormente determinam diretamente o valor de todos os parâmetros existentes. Contudo, mais tarde foram desenvolvidas técnicas sofisticadas que garantiam uma forma mais rápida e menos dispendiosa de encontrar indiretamente os parâmetros. Surgiram então as séries *Even-Tempered* e *Well-Tempered*.

A primeira série *Even-Tempered* foi proposta em (RUEDENBERG *et al*, 1972) e mais tarde em (SCHMIDT & RUEDENBERG, 1979). Nessa série, os expoentes são relacionados pela série geométrica dada pela equação (3-9).

$$\zeta_k = \alpha\beta^k, k = 1, \dots, N_l \quad (3-9)$$

$N_l$  é o número de primitivas gaussianas na simetria  $l$ , e os parâmetros  $\alpha_l$ ,  $\beta_l$  são parâmetros reais arbitrários (BARDO & RUEDENBERG, 1973).

Mais tarde, as séries *Even-Tempered* foram modificadas em (HUZINAGA *et al*, 1985). Nesse trabalho essas séries ganharam mais um parâmetro e tomaram

a forma da equação (3-10), sendo nomeadas de séries *Well-Tempered*. Utilizando essas séries, (HUZINAGA & KLOBUKOWSKI, 1988) desenvolveram conjuntos de base para os átomos do Li ao Hg.

$$\zeta_k = \alpha\beta^{N-k+1} \left[ 1 + \lambda \left( \frac{N-k+1}{N} \right)^\delta \right], k = 1, \dots, N_l \quad (3-10)$$

Foram também introduzidas fórmulas recursivas para gerar coeficientes para as primitivas gaussianas em (KLOBUKOWSKI, 1993). As fórmulas são dadas pelas equações (3-11) e (3-12).

$$\zeta_1 = P_1 P_2 \quad (3-11)$$

$$\zeta_k = \zeta_{k-1} P_2 [1 + P_3(k-1) + P_3(k-1)^2], k = 2, \dots, N \quad (3-12)$$

Alguns métodos recursivos de se gerar coeficientes ganharam destaque; um deles foi o método da coordenada geradora (MCG) introduzido por Griffin, Hill e Wheeler nos anos 50 (HILL, 1953) (GRIFFIN, 1957) e modificado em (MOHALLEM, 1986). Nesse trabalho é mostrado que o aproveitamento do método MCG é alcançado quando uma discretização é aplicada com o propósito de se obter uma melhor integração numérica da equação GHW (MOHALLEM, 1986). O método MCG foi amplamente aplicado em física nuclear (WONG, 1975) e posteriormente em sistemas atômicos e moleculares (CHATTOPADHYAYFS, 1981) (LATHOUWERS, 1978).

Em 1986, o método da coordenada geradora Hartree-Fock (GCHF) foi introduzido como uma nova técnica para gerar funções de base do tipo gaussianas (GTF) e do tipo Slater (STF). O método GCHF foi usado com sucesso na geração de funções de base gaussianas universais (UGBSs) (JORGE, 1997) e funções de base gaussianas adaptadas (AGBSs) (JORGE, 1999) para átomos leves e pesados.

No método GCHF, as funções de um elétron são escolhidas como uma transformada integral (MOHALLEM, 1986), isto é:

$$\psi_i = \int \phi_i(\mathbf{1}, \alpha) f_i(\alpha), i = 1, \dots, n \quad (3-13)$$

onde  $\phi_i$  são as funções geradoras (devem ser STF, GTF ou outro tipo de função),  $f_i$  são as funções de peso e  $\alpha$  é a coordenada geradora. A solução é realizada através da escolha apropriada de um conjunto discreto de pontos no espaço da coordenada geradora, representada por:

$$\Omega_k = \Omega_{min} + (k-1)\Delta\Omega, k = 1, \dots, N \quad (3-14)$$

Portanto, o método GCHF original usa apenas uma sequência aritmética de pontos igualmente espaçados, como mostra a equação 5-2, para gerar as funções de base (MOHALLEM, 1986). O conjunto de base gerado por este método é conhecido como SOGBs (JORGE *et al*, 1999).

Com o intuito de aprimorar o método descrito anteriormente, foi proposto em (JORGE *et al*, 1999) o uso de duas ou três sequências aritméticas.

Nesta nova alternativa, as coordenadas geradoras são discretizadas, para cada simetria, em duas ou em três sequências aritméticas independentes, dadas pela Equação (3-15). O número de parâmetros a ser otimizado para as Equações (3-15) e (3-16) é, respectivamente, duas e três vezes maior do que o GCHF original.

$$\Omega_k = \begin{cases} \Omega_{min} + (k - 1)\Delta\Omega, k = 1, \dots, M \\ \Omega_{min}^1 + (k - 1)\Delta\Omega, k = M + 1, \dots, N \end{cases} \quad (3-16)$$

$$\Omega_k = \begin{cases} \Omega_{min} + (k - 1)\Delta\Omega, k = 1, \dots, J \\ \Omega_{min}^1 + (k - 1)\Delta\Omega^1, k = J + 1, \dots, M \\ \Omega_{min}^2 + (k - 1)\Delta\Omega^2, k = M + 1, \dots, N \end{cases} \quad (3-17)$$

onde,

$$\Omega_k = \ln \alpha_k / A, A > 1. \quad (3-18)$$

Formas não-recursivas de se gerar os parâmetros foram também estudadas em (KLOBUKOWSKI, 1990) e em (KLOBUKOWSKI, 1993). São introduzidas nestes trabalhos as equações (3-19) e (3-20), ambas tiveram boa aceitação quando a questão foi gerar boas parametrizações para conjuntos de base.

$$\zeta_k = \sum_{i=0}^n P_i k^{i-1} \quad (3-19)$$

$$\zeta_k = \frac{P_1 + P_2k + P_3k^2}{1 + P_4k + P_5k^2}$$

(3-20)