



Fundação Universidade Federal do ABC
Pró reitoria de pesquisa

Av. dos Estados, 5001, Santa Terezinha, Santo André/SP, CEP 09210-580
Bloco L, 3º Andar, Fone (11) 3356-7617
iniciacao@ufabc.edu.br

Relatório Final de Iniciação Científica referente ao Edital: 03/2023

Nome do aluno: Felipe Fernandes Gomes da Silva Costa

Assinatura do aluno:

Nome do Orientador: Yuri Alexandre Aoto

Assinatura do orientador:

Título do projeto: Otimização de Geometria Molecular

Palavras-chave do projeto: Superfície de Energia Potencial, Otimização Geométrica, Química Computacional, Método de Newton, Método das Secantes.

Área do conhecimento do projeto: Matemática Aplicada

Bolsista: Não

Santo André, SP
30 de setembro de 2024

Sumário

1	Resumo	3
2	Introdução	3
3	Definições	4
3.1	Otimização	4
3.2	Iterações	4
4	Fundamentação Teórica	5
4.1	Método de Newton	5
4.2	Método da Secante	6
4.3	Método de Newton Multidimensional	6
4.4	Reação $F + H_2O \longrightarrow FH + HO$	7
4.5	Módulo Fortran para Função SEP	8
5	Metodologia	8
5.1	Parâmetros de Referência	8
5.1.1	Otimização com Método de Newton	9
5.1.2	Evitando Casos de Matriz Singular	9
5.2	Funcionamento CBPD	11
5.2.1	Interpretação do Método	12
5.3	Avaliando Eficiência do Algoritmo	12
5.3.1	Variando uma Coordenada	12
5.3.2	Cenários Aleatórios de Convergência	14
6	Resultados	14
6.1	Variando Uma Variável	14
6.2	Cenários Aleatórios de Convergência	15
7	Conclusões	15

1 Resumo

A otimização de geometria molecular é um tópico essencial da química computacional, em que utilizando funções de Superfície de Energia Potencial (SEP) é possível prever configurações estáveis de uma molécula ou estados de transição. O objetivo deste projeto foi desenvolver um algoritmo para otimização de funções SEP baseado no método de Newton, mantendo uma taxa de convergência similar e com menor custo computacional. O algoritmo proposto, CBPD (Convergence Based on Partial Derivatives), foi aplicado à reação $F + H_2O \longrightarrow FH + HO$, permitindo avaliar sua eficácia em diferentes cenários de otimização. Comparado ao método de Newton, o CBPD mostrou boas taxas de convergência em cenários específicos, apesar de exigir um número maior de iterações cada etapa de iteração possui um menor custo computacional. Os resultados indicam que o CBPD pode ser uma alternativa promissora para problemas de otimização em química computacional de funções SEP em um espaço de alta dimensionalidade.

2 Introdução

A otimização geométrica de estruturas moleculares é um ramo de pesquisa da química computacional que busca por meio de algoritmos identificar conformações moleculares que são estáveis, ou que sejam uma configuração de uma etapa de uma reação química. É de interesse compreender essas conformações não apenas para o entendimento dos mecanismos das reações, mas também para aplicações práticas, como o desenvolvimento de novos fármacos.

O conceito de Superfície de Energia Potencial (SEP - do inglês, *Potential Energy Surface*) descreve que para cada configuração de uma molécula, i.e, o tamanho das ligações e ângulos em que os átomos estão arranjados, existe um valor de energia associado. Dado que uma SEP pode estar associada a uma função, localizar os pontos críticos dessa função usualmente representa identificar as configurações estáveis de uma molécula ou estados de transição, que estão diretamente relacionados com mínimos locais e pontos de sela respectivamente [4].

Funções que representam SEP, por conta de suas complexidades, não costumam possuir expressões analíticas. Dessa maneira, os valores de suas derivadas, que são frequentemente utilizados em métodos de otimização, passam a ser obtidos exclusivamente por aproximações. Diante disso, é relevante que o método de otimização implementado possua uma maneira eficiente de fazer o cálculo das derivadas.

O objetivo de estudo do projeto é a construção de um algoritmo baseado no Método de Newton e Método da Secante, que permita realizar a otimização de funções SEP. O

algoritmo trata o processo de otimização de uma função multidimensional considerando apenas a diagonal de sua matriz hessiana, como será apresentado na seção 5.2. No caso, é feito a otimização de cada derivada parcial da função multidimensional utilizando o Método de Newton e aproximando os valores das derivadas parciais utilizando o Método da Secante.

Para a validação do algoritmo desenvolvido será utilizado a função SEP que descreve a reação entre fluor e água ($F + H_2O \longrightarrow FH + HO$) [2]. Apesar desta função possuir uma expressão analítica, o trabalho visa desenvolver um método genérico de convergência de funções PES. O algoritmo desenvolvido denominado CBPD (*Convergence Based in Partial Derivatives*) será comparado com o Método de Newton em diferentes cenários de convergência, avaliando a taxa de convergência e a quantidade de iterações necessárias para convergir.

A organização do restante deste resumo é descrita a seguir. A Seção 3 apresenta termos fundamentais em métodos de otimização. A Seção 4 revisa os métodos de convergência que serão utilizados como base, assim como descreve sobre a reação de estudo. A metodologia utilizada nesse projeto é descrita na Seção 5. Por fim os resultados e conclusões são apresentados nas Seções 6 e 7 respectivamente.

3 Definições

3.1 Otimização

O processo de otimização do contexto desse trabalho consiste em localizar os mínimos locais da função que descreve a superfície de energia potencial que está sendo estudada. Nesse processo, fornecendo uma geometria inicial da molécula em estudo, ou seja, um valor inicial para os argumentos da função que descreve a SEP, será retornado a geometria em que a SEP tem valor mínimo local. O processo é feito por iterações até que o algoritmo de otimização consiga convergir (otimizar) ou chegue no limite de iterações definidas.

3.2 Iterações

Uma iteração, nesse contexto, consiste em cada etapa no processo de otimização. Inicialmente temos ponto do domínio da função, é realizado então a etapa de convergência, que é justamente o enfoque do projeto, que consiste em realizar um processo matemático para definir um próximo ponto do domínio que idealmente deve ser mais próximo do mínimo local da função. Caso o novo ponto definido esteja próximo o suficiente do mínimo local, é dito que o algoritmo convergiu e o processo de otimização é finalizado. A definição

se o ponto está próximo o suficiente é feita com o uso de norma, no projeto atual está sendo usado a norma euclidiana. Para cada novo ponto é calculada a norma do gradiente da função, caso esse valor seja igual ou menor a um valor de tolerância é definido que algoritmo convergiu. O valor de tolerância é um parâmetro que define o quão próximo o ponto deve estar do mínimo local para ser considerado que o ponto está no ponto mínimo. Quanto menos iterações forem feitas e quanto menor o custo computacional envolvido em cada iteração, é dito que o algoritmo é mais eficiente.

4 Fundamentação Teórica

O processo de otimização pode ser realizado com diferentes métodos, que nesse contexto possuem o mesmo objetivo, localizar as raízes de uma função, seja essa função fornecida analiticamente ou numericamente. Todo método existente possui seus pontos positivos e negativos dentre os demais, seja por sua eficiência, facilidade de implementação ou custo computacional relacionado a cada etapa da iteração. A pesquisa está se embasando principalmente sobre o Método de Newton, que satisfazendo seus critérios de convergência [6] é classificado como estável, ou seja, a cada iteração é obtido um novo ponto mais próximo da região de convergência que possui um erro relativo menor quando comparado com a etapa anterior.

4.1 Método de Newton

No Método de Newton, dado um ponto inicial, utilizando a derivada da função em estudo, se obtém um próximo ponto mais próximo da região de convergência. Nesse processo, iniciamos de um ponto inicial arbitrário da função e traçamos uma reta tangente da função no ponto, ou seja, verificamos o valor da derivada no ponto e utilizamos para definir uma função afim $g : \mathbb{R} \rightarrow \mathbb{R}$ com $g(x) = ax + b$, $a, b \in \mathbb{R}$. Com a função afim definida no ponto, identificamos a raiz da função, ou seja, definimos o ponto em que $g(x) = 0$.

Podemos calcular o ponto x_n em que $g(x) = 0$ da seguinte maneira:

$$x_n = x_{n-1} - \frac{f(x)}{f'(x)} \quad (1)$$

Realizamos novamente o processo anterior, agora para o ponto $f(x_n)$. Para cada novo ponto x_k obtido aproximamos cada vez mais do ponto em que $f(x_k)$ é igual a zero.

4.2 Método da Secante

O Método da Secante utiliza dois pontos da função para calcular o próximo ponto da iteração com o objetivo que seja mais próximo da raiz da função. Cada iteração é calculada com base na reta formada pelos dois pontos anteriores, que será secante ao gráfico da função. Dessa forma, o ponto x_n é definido:

$$x_n = x_{n-1} - f(x_{n-1}) \frac{x_{n-1} - x_{n-2}}{f(x_{n-1}) - f(x_{n-2})}. \quad (2)$$

Trazendo enfoque na etapa de iteração, mais precisamente no termo relativo a reta secante

$$\frac{x_{n-1} - x_{n-2}}{f(x_{n-1}) - f(x_{n-2})}, \quad (3)$$

se definirmos $x_{n-1} = x_{n-2} + h$ com $h \in \mathbb{R}$ temos

$$\frac{x_{n-2} + h - x_{n-2}}{f(x_{n-2} + h) - f(x_{n-2})} = \frac{h}{f(x_{n-2} + h) - f(x_{n-2})}. \quad (4)$$

Podemos tomar o inverso dessa expressão, e, para casos em quem que a distância dentre os dois pontos seja suficientemente pequena, ou seja, caso o valor de h tenda a 0, o termo representado passa a ser o inverso de uma aproximação da derivada no ponto.

$$\left(\frac{f(x_{n-2} + h) - f(x_{n-2})}{h} \right)^{-1} \approx \left(\lim_{h \rightarrow 0} \frac{f(x_{n-2} + h) - f(x_{n-2})}{h} \right)^{-1} = \left(\frac{df}{dx}(x_{n-2}) \right)^{-1} \quad (5)$$

Além disso, notemos que a função f é diferenciável, o Teorema do Valor Médio [5] afirma que existe uma reta tangente entre os dois pontos x_{n-1} e x_{n-2} que o seu valor é exatamente o valor da reta secante calculada.

4.3 Método de Newton Multidimensional

O Método de Newton Multidimensional consiste em uma generalização do Método de Newton porém para casos que possam envolver funções multidimensionais. Sejam uma função $F : \mathbb{R}^k \rightarrow \mathbb{R}^k$ com $k \in \mathbb{N}$ e $\mathbf{x}_n \in \mathbb{R}^k$ referente ao enésimo ponto do processo de otimização. Um novo ponto \mathbf{x}_{n+1} é definido por:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - J_F(\mathbf{x}_n)^{-1} F(\mathbf{x}_n), \quad (6)$$

sendo $J_F(\mathbf{x}_n)^{-1}$ a matriz inversa $k \times k$ do Jacobiano da função F .

É importante se atentar nesse método sobre quais condições a etapa de iteração pode

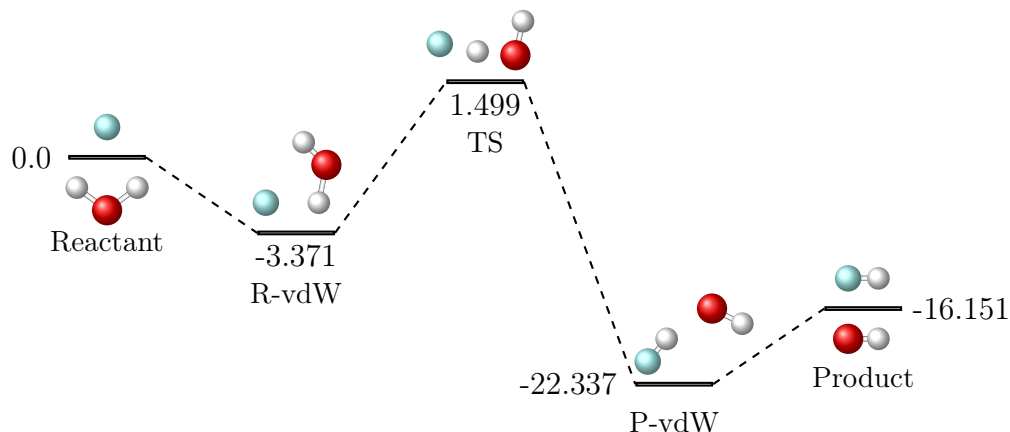


Figura 1: Representação gráfica do perfil SEP da reação $F + H_2O \rightarrow FH + HO$.

performar. Nesse caso é sempre necessário verificar se a matriz $J_F(\mathbf{x}_n)$ é inversível, ou seja, $\det(J_F(\mathbf{x}_n)) \neq 0$, pois caso contrário não é possível dar sequência no processo de otimização.

4.4 Reação $F + H_2O \rightarrow FH + HO$

Nessa pesquisa será estudada a reação de $F + H_2O \rightarrow FH + HO$ que passa por 5 pontos estacionários. Cada ponto estacionário possui uma conformação específica que permitem que os reagentes interajam. No caso em estudo são:

- Reagentes
- R-vdW
- TS
- P-vdW
- Produtos

Cada ponto estacionário é caracterizado por uma geometria específica que possui um valor de energia associado, que é uma consequência da conformação geométrica dos átomos e suas interações. Esses átomos ficam configurados de maneira que proporcionam a energia mínima para que cada ponto estacionário da reação ocorra. Possuindo a função que descreve a energia de cada conformação de uma dada reação, é possível determinar a configuração geométrica ótima para cada ponto estacionário localizando o mínimo local da função.

4.5 Módulo Fortran para Função SEP

Nessa pesquisa, será utilizado um módulo implementado em Fortran [3], o qual possui uma interface em Python. Essa interface permite a inserção de configurações geométricas da reação $F + H_2O \longrightarrow FH + HO$ como entrada e, como resultado, retorna o valor da energia associada a essa configuração. A função em Python recebe como parâmetro uma lista de tamanho 6, $\mathbf{x}_n \in \mathbb{R}^6$, que representam cada configuração da reação e recebe como retorno o valor de energia associado em kcal/mol.

Tabela 1: Relação entre variáveis esperadas pela função SEP e quais coordenadas representam na reação $F + H_2O \longrightarrow FH + HO$.

Variável	Coordenada
x_1	Distância H–O
x_2	Distância O–H'
x_3	Distância H'–F
x_4	Ângulo HOH'
x_5	Ângulo OH'F
x_6	Ângulo Diedro HOH'F

5 Metodologia

A metodologia adotada para desenvolver um algoritmo de otimização consiste em definir os parâmetros de referência que buscamos atingir, desenvolver o algoritmo de otimização e aplicá-lo em cada ponto estacionário da reação buscando obter os mesmos resultados de referência. Todos os códigos gerados e utilizados para obter os resultados desse projeto estão disponíveis no repositório do GitHub [1].

5.1 Parâmetros de Referência

O artigo [2] apresenta as conformações otimizadas para cada etapa da reação. Porém existe uma diferença do ponto mínimo apresentado no artigo com o obtido no módulo fortran [3] com interface em python, que retorna o valor de energia associado a uma determinada conformação geométrica. Dessa forma, é necessário reajustar as configurações de referência com base no módulo fortran.

O processo de ajuste dos parâmetros de referência foi feito com diferentes métodos de otimização e foi escolhido o que apresentou as melhores taxas de convergência. Os métodos usados foram:

- BFGS - Presente na biblioteca SciPy [7]

- CG - Presente na biblioteca SciPy [7]
- Newton CG - Presente na biblioteca SciPy [7]
- Método de Newton

Dos métodos utilizados o que se apresentou mais eficaz foi o Método de Newton, que conseguiu convergir em mais casos mesmo com variações na conformação inicial.

5.1.1 Otimização com Método de Newton

Como apresentado na equação (6), cada etapa de iteração é feita calculando o inverso do Jacobiano da função $G : \mathbb{R}^k \rightarrow \mathbb{R}^k, k \in \mathbb{N}$. Porém no cenário de estudo dessa pesquisa temos uma função $F : \mathbb{R}^k \rightarrow \mathbb{R}, k \in \mathbb{N}$, e para buscar o seu mínimo local, devemos localizar o ponto do domínio em que a norma do gradiente da função (∇F) seja zero. Dessa maneira, a etapa de iteração é calculada

$$\mathbf{x}_{n+1} = \mathbf{x}_n - H_F(\mathbf{x}_n)^{-1} \nabla F(\mathbf{x}_n), \quad (7)$$

sendo $H_F(\mathbf{x}_n)^{-1}$ a matriz inversa $k \times k$ da hessiana da função F e ∇F o vetor gradiente da função F . Note que F e \mathbf{x}_n são, respectivamente, a função SEP e as configurações da geometria da moléculas conforme descrito na seção 4.5.

O vetor gradiente e a matriz hessiana utilizam das primeiras e segundas derivadas parciais respectivamente para serem calculadas. Por estarmos tratando de funções numéricas, não é possível obter uma expressão analítica das derivadas parciais. Nesse caso, está sendo feito o cálculo da derivada parcial numericamente utilizando a expressão

$$\frac{\partial F}{\partial x_i} \approx \frac{F(\mathbf{x} + h\mathbf{e}_i) - F(\mathbf{x} - h\mathbf{e}_i)}{2h}, \quad (8)$$

onde \mathbf{e}_i o iésimo vetor da base canônica de \mathbb{R}^k e o valor de $h = 10^{-6}$. Dessa maneira é obtido uma média das derivadas parciais que tenderiam para direita e para a esquerda.

5.1.2 Evitando Casos de Matriz Singular

Esse método tem o processo de iteração interrompido caso a matriz hessiana possuir determinante igual a zero (matriz singular), que são os casos em que a matriz não possui inversa. Notamos que no nosso problema essa situação ocorre quando ao menos um valor do gradiente da função é localmente constante igual a zero, ou seja, a função é localmente constante ao variar uma das coordenada. Isso ocorre, pois no momento de calcular a matriz hessiana haverá ao menos uma linha ou coluna com todos os valores zerados, consequentemente o determinante da matriz hessiana também será zero.

Os parâmetros dessa função descrevem a distância de ligação entre os átomos e os ângulos formados entre eles. Dependendo para qual ponto estacionário a geometria está sendo otimizada, alguns parâmetros não interferem no valor da função que retorna a energia associada a determinada configuração. Nesses casos, durante o processo de otimização é de interesse remover esses parâmetros na etapa de iteração, pois diminui a chance de se obter a matriz hessiana singular. Além disso, pelo fato de se estar calculando menos derivadas o processo de otimização fica mais rápido.

Usando como exemplo a etapa o primeiro ponto estacionário $F + H_2O$, apenas 3 coordenadas são de interesse para se otimizar, como pode ser visualizado na tabela 2.

Tabela 2: Coordenadas de interesse para otimização do ponto estacionário $F + H_2O$.

Ponto Estacionário		H – O (x_1)	O – H' (x_2)	H' – F (x_3)	HOH' (x_4)	OH'F (x_5)	HOH'F (x_6)
$F + H_2O$	Otimizar	✓	✓		✓		

Dessa forma, reduzimos a complexidade do cálculo da matriz hessiana e do vetor gradiente, que deixam de serem calculadas por:

$$H_F = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} & \frac{\partial^2 f}{\partial x_1 \partial x_4} & \frac{\partial^2 f}{\partial x_1 \partial x_5} & \frac{\partial^2 f}{\partial x_1 \partial x_6} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} & \frac{\partial^2 f}{\partial x_2 \partial x_4} & \frac{\partial^2 f}{\partial x_2 \partial x_5} & \frac{\partial^2 f}{\partial x_2 \partial x_6} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} & \frac{\partial^2 f}{\partial x_3 \partial x_4} & \frac{\partial^2 f}{\partial x_3 \partial x_5} & \frac{\partial^2 f}{\partial x_3 \partial x_6} \\ \frac{\partial^2 f}{\partial x_4 \partial x_1} & \frac{\partial^2 f}{\partial x_4 \partial x_2} & \frac{\partial^2 f}{\partial x_4 \partial x_3} & \frac{\partial^2 f}{\partial x_4^2} & \frac{\partial^2 f}{\partial x_4 \partial x_5} & \frac{\partial^2 f}{\partial x_4 \partial x_6} \\ \frac{\partial^2 f}{\partial x_5 \partial x_1} & \frac{\partial^2 f}{\partial x_5 \partial x_2} & \frac{\partial^2 f}{\partial x_5 \partial x_3} & \frac{\partial^2 f}{\partial x_5 \partial x_4} & \frac{\partial^2 f}{\partial x_5^2} & \frac{\partial^2 f}{\partial x_5 \partial x_6} \\ \frac{\partial^2 f}{\partial x_6 \partial x_1} & \frac{\partial^2 f}{\partial x_6 \partial x_2} & \frac{\partial^2 f}{\partial x_6 \partial x_3} & \frac{\partial^2 f}{\partial x_6 \partial x_4} & \frac{\partial^2 f}{\partial x_6 \partial x_5} & \frac{\partial^2 f}{\partial x_6^2} \end{bmatrix}, \nabla F = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \\ \frac{\partial f}{\partial x_4} \\ \frac{\partial f}{\partial x_5} \\ \frac{\partial f}{\partial x_6} \end{bmatrix}, \quad (9)$$

passando a serem calculadas por:

$$H_F = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_4} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_4} \\ \frac{\partial^2 f}{\partial x_4 \partial x_1} & \frac{\partial^2 f}{\partial x_4 \partial x_2} & \frac{\partial^2 f}{\partial x_4^2} \end{bmatrix}, \nabla F = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_4} \end{bmatrix}. \quad (10)$$

Aplicando o Método de Newton com base nos valores de referência dos pontos estacionários do artigo [2], foi definido as conformações a serem utilizadas como parâmetro para o presente trabalho e podem ser verificadas na tabela 3.

Tabela 3: Geometrias nos Pontos Estacionários. As coordenadas relevantes referem-se as interações que mais influenciam no valor energético do ponto estacionário em estudo, essas coordenadas são as que serão otimizadas. Os valores associados as demais coordenadas são escolhidos de forma para não influenciar o processo de otimização.

Ponto Estacionário		H – O (x_1)	O – H' (x_2)	H' – F (x_3)	HOH' (x_4)	OH'F (x_5)	HOH'F (x_6)
F + H ₂ O	Vars. Relevantes	✓	✓		✓		
	Valor	0.9609 Å	0.9609 Å	20 Å	104.1477°	300°	300°
R – vdW (F...H ₂ O)	Vars. Relevantes	✓	✓	✓	✓	✓	✓
	Valor	0.9641 Å	0.9641 Å	2.2981 Å	104.3416°	67.1171°	–87.5402°
TS	Vars. Relevantes	✓	✓	✓	✓	✓	✓
	Valor	0.9693 Å	1.0250 Å	1.3536 Å	103.2989°	118.2898°	68.3691°
P – vdW (HO...HF)	Vars. Relevantes	✓	✓	✓	✓	✓	✓
	Valor	0.9728 Å	1.7700 Å	0.9354 Å	108.6504°	173.6165°	–0.0737°
HO + HF	Vars. Relevantes	✓		✓			
	Valor	0.9728 Å	20 Å	0.9223 Å	300°	300°	300°

5.2 Funcionamento CBPD

O método CBPD (Convergence Based in Partial Derivatives) de otimização, desenvolvido nessa pesquisa, se baseia em otimizar uma função localizando a raiz de cada uma das suas derivadas parciais individualmente. Dada uma função $F : \mathbb{R}^k \rightarrow \mathbb{R}, k \in \mathbb{N}$, o método busca identificar o ponto em que a norma euclidiana das derivadas parciais seja menor que um valor de tolerância, sendo um valor próximo de zero

$$\sqrt{\sum_{i=1}^k \left| \frac{\partial F}{\partial x_i} \right|^2} < \text{tolerance} \approx 0. \quad (11)$$

Nesse método, cada etapa da iteração é calculada

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \text{diag} \left(\frac{\partial^2 F}{\partial (x_n)^2} \right)^{-1} \nabla F(\mathbf{x}_n), \quad (12)$$

sendo diag a matriz diagonal da matriz hessiana da função F . Cada coordenada $(x_{n+1})_i$ do vetor \mathbf{x}_{n+1} pode ser calculada por

$$(x_{n+1})_i = (x_n)_i - \left(\frac{\partial^2 F}{\partial (x_n)_i^2} \right)^{-1} \frac{\partial F}{\partial (x_n)_i}, i \in \{1, \dots, k\}, \quad (13)$$

em que, utilizando da notação definida em (8), definimos

$$\begin{aligned}\frac{\partial F}{\partial(x_n)_i} &\approx \frac{F(\mathbf{x}_n + h\mathbf{e}_i) - F(\mathbf{x}_n - h\mathbf{e}_i)}{2h} \\ \frac{\partial^2 F}{\partial(x_n)_i^2} &\approx \frac{\frac{\partial F(\mathbf{x}_n + h\mathbf{e}_i)}{\partial(x_n)_i} - \frac{\partial F(\mathbf{x}_n - h\mathbf{e}_i)}{\partial(x_n)_i}}{2h}.\end{aligned}\tag{14}$$

5.2.1 Interpretação do Método

O método se baseia em olhar para o problema de otimização tratando cada variável da função a ser otimizada de maneira independente. Tomamos cada derivada parcial $\frac{\partial F}{\partial x_i}$ da função $F \in \mathbb{R}^k$ e aplicamos o Método de Newton (4.1) em cada uma dessas derivadas parciais, usando as aproximações para a primeira e segunda derivada parcial definidas em (14).

Tomamos como exemplo a função $F(x, y) = 1.5x^2 + 1.2y^2 - 0.25x^4 - 0.3y^4$ com o ponto inicial $\mathbf{P}_0 = (x_0, y_0) = (1, 0.8)$ que possui como derivadas parciais

$$\begin{aligned}\frac{\partial F}{\partial x} &= 3x - x^3 \\ \frac{\partial F}{\partial y} &= 2.4y - 1.2y^3.\end{aligned}\tag{15}$$

Uma etapa de iteração para calcular o ponto $\mathbf{P}_1 = (x_1, y_1)$ é feita calculando

$$\begin{aligned}x_1 &= x_0 - \frac{1}{3x_0 - x_0^3}F(x_0, y_0) = 0.05 \\ y_1 &= y_0 - \frac{1}{2.4y_0 - 1.2y_0^3}F(x_0, y_0) \approx -0.65,\end{aligned}\tag{16}$$

obtendo $\mathbf{P}_1 = (x_1, y_1) = (0.05, -0.65)$. Geometricamente esse processo é interpretado como na figura 2.

O procedimento é repetido até atingir o critério de convergência (11) ou até atingir um valor máximo arbitrário de iterações.

5.3 Avaliando Eficiência do Algoritmo

5.3.1 Variando uma Coordenada

A primeira estratégia que será utilizada é, para cada ponto estacionário, como que o algoritmo se comporta quando apenas uma coordenada se distancia da conformação ideal. Nessa abordagem todas as coordenadas ficam em seu valor ótimo com exceção da

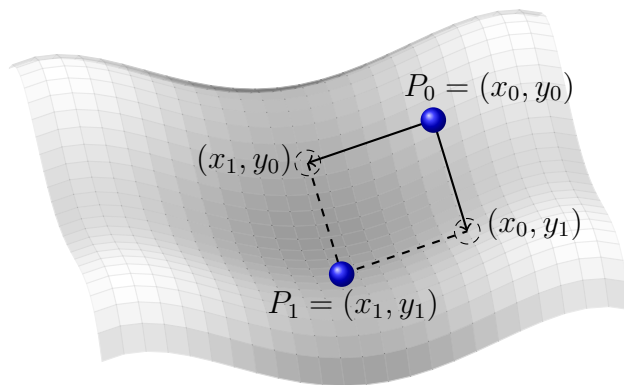


Figura 2: Representação geométrica de uma etapa do método de convergência CBPD.

coordenda que está variando. Dessa maneira estaremos trabalhando com um problema em apenas uma dimensão.

Para cada ponto estacionário da reação será analisado uma coordenada relevante de cada vez, variando ela em -25% , -20% , -15% , 10% , -5% , 0% , 5% , 10% , 15% , 20% , 25% do seu valor otimizado.

Exemplificando para o caso $F + H_2O$ haverão 11 cenários de tentativa de convergência para a primeira coordenada (Ligação $H + O$) como apresentados na tabela 4. O mesmo processo será feito para todas as demais coordenadas relevantes.

Tabela 4: Cenários de convergência do método de variação de uma coordenada para a coordenada $H-O$ do ponto estacionário $F + H_2O$

Caso	H – O (x_1)	O – H' (x_2)	H' – F (x_3)	HOH' (x_4)	OH'F (x_5)	HOH'F (x_6)
1	0.7207 Å	0.9609 Å	20 Å	104.1477°	300°	300°
2	0.7687 Å	0.9609 Å	20 Å	104.1477°	300°	300°
3	0.8168 Å	0.9609 Å	20 Å	104.1477°	300°	300°
4	0.8648 Å	0.9609 Å	20 Å	104.1477°	300°	300°
5	0.9128 Å	0.9609 Å	20 Å	104.1477°	300°	300°
6	0.9609 Å	0.9609 Å	20 Å	104.1477°	300°	300°
7	1.0089 Å	0.9609 Å	20 Å	104.1477°	300°	300°
8	1.0570 Å	0.9609 Å	20 Å	104.1477°	300°	300°
9	1.1050 Å	0.9609 Å	20 Å	104.1477°	300°	300°
10	1.1531 Å	0.9609 Å	20 Å	104.1477°	300°	300°
11	1.2011 Å	0.9609 Å	20 Å	104.1477°	300°	300°

5.3.2 Cenários Aleatórios de Convergência

A segunda métrica que será utilizada busca explorar a convergência quando múltiplas coordenadas não estão em sua configuração ótima. Para isso, para cada etapa da reação química que possui a sua configuração otimizada, serão geradas 100 configurações distorciadas, nas quais a variação máxima para as ligações serão de ± 0.3 Å e para ângulos serão de $\pm 10^\circ$. Os valores mínimos e máximos para cada estado estacionário podem ser verificados na tabela 5.

Tabela 5: Valores mínimos e máximos que podem ser gerados nos cenários aleatórios de convergência para cada ponto estacionário da reação $F + H_2O \longrightarrow FH + HO$

Ponto Estacionário		H – O (x_1)	O – H' (x_2)	H' – F (x_3)	HOH' (x_4)	OH'F (x_5)	HOH'F (x_6)
F + H ₂ O	Valor Mín.	0.6609 Å	0.6609 Å	19.7 Å	94.1477°	290°	290°
	Valor Máx.	1.2609 Å	1.2609 Å	20.3 Å	114.1477°	310°	310°
R – vdW (F...H ₂ O)	Valor Mín.	0.6641 Å	0.6641 Å	1.9981 Å	94.3416°	57.1171°	–97.5402°
	Valor Máx.	1.2641 Å	1.2641 Å	2.5981 Å	114.3416°	77.1171°	–77.5402°
TS	Valor Mín.	0.6693 Å	0.7270 Å	1.0536 Å	93.2989°	108.2898°	58.3691°
	Valor Máx.	1.2693 Å	1.3250 Å	1.6536 Å	113.2989°	128.2898°	78.3691°
P – vdW (HO...HF)	Valor Mín.	0.6728 Å	1.4700 Å	0.6354 Å	98.6504°	163.6165°	–10.0737°
	Valor Máx.	1.2728 Å	2.0700 Å	1.2354 Å	118.6504°	183.6165°	9.9263°
HO + HF	Valor Mín.	0.6728 Å	19.7 Å	0.6223 Å	290°	290°	290°
	Valor Máx.	1.2728 Å	20.3 Å	1.2223 Å	310°	310°	310°

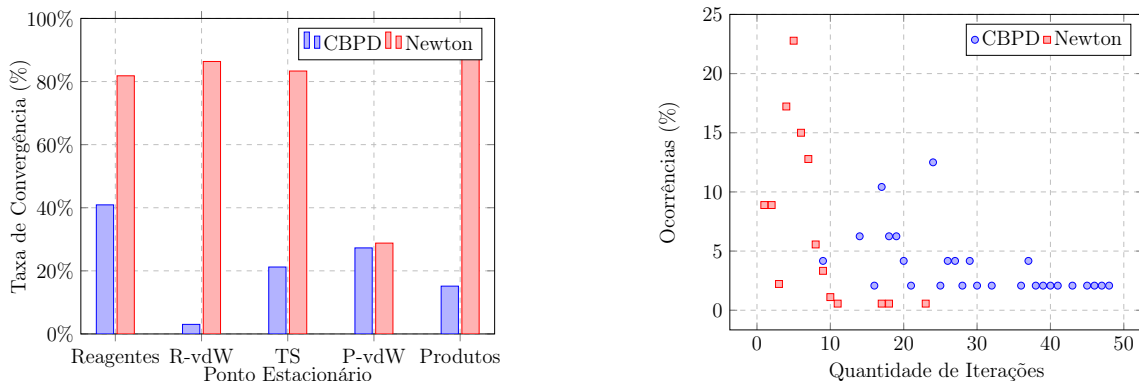
6 Resultados

Os resultados obtidos levam em consideração a quantidade de casos que houveram convergência, e dentre esses casos, a quantidade de iterações necessárias. As informações são apresentadas de forma comparativa com o método Newton nos mesmos cenários de convergência. Todos os testes de convergência possuem um limite de 50 iterações, ao ultrapassá-lo é assumido que o algoritmo não convergiu.

6.1 Variando Uma Variável

Para os cenários onde uma variável de cada vez era afastado de seu estado otimizado, foi feita uma análise da taxa de sucesso de convergência por ponto estacionário como apresentado no gráfico 3a. Para os casos de sucesso de convergência, um gráfico de

dispersão foi montado apresentando o percentual da quantidade de casos de convergência por número de iterações necessárias conforme apresentado no gráfico 3b.



(a) Taxa de sucesso de convergência para os métodos iterativos CBPD e Newton para cada ponto estacionário da reação $F + H_2O \longrightarrow FH + HO$ nos casos de variação de uma variável.

(b) Percentual da quantidade de iterações necessárias para convergir para os métodos iterativos CBPD e Newton para a reação $F + H_2O \longrightarrow FH + HO$ nos casos de variação de uma variável.

Figura 3: Resultados dos casos de variação de uma variável.

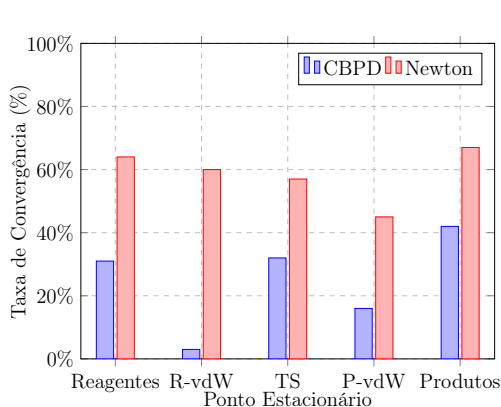
6.2 Cenários Aleatórios de Convergência

Para os cenários de convergência nos quais todas as variáveis são aleatorizadas, foi feita uma análise similar à descrita na seção 6.1. Uma análise da taxa de convergência por ponto estacionário é apresentada no gráfico 4a, enquanto para os casos de sucesso convergência, um gráfico de dispersão foi montado apresentando o percentual da quantidade de casos de convergência por número de interações conforme apresentado no gráfico 4b.

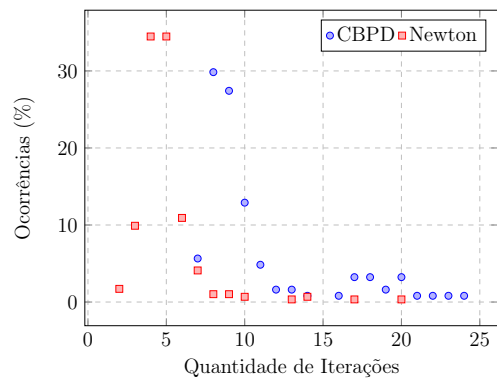
Avaliando as taxas de convergências dos métodos CBPD e Newton é notório que o método de Newton apresenta melhores resultados de convergência na maioria dos pontos estacionários da reação com exceção do $P - vdW$. Em termos de quantidades de iterações necessárias para convergência, o Método de Newton apresenta uma maior consistência, sendo necessário em torno de 5 a 7 iterações. Já no Método CBPD a quantidade de iterações necessárias possui um pico entre 11 e 12 iterações.

7 Conclusões

A presente pesquisa permitiu explorar um novo método iterativo aplicado em um cenário real, de maneira comparativa com o método de Newton e permitindo evidenciar resultados positivos e negativos em diferentes cenários de convergência, avaliando



(a) Taxa de sucesso de convergência para os métodos iterativos CBPD e Newton para cada ponto estacionário da reação $F + H_2O \longrightarrow FH + HO$ nos cenários aleatórios de convergência.



(b) Percentual da quantidade de iterações necessárias para convergir para os métodos iterativos CBPD e Newton para a reação $F + H_2O \longrightarrow FH + HO$ nos cenários aleatórios de convergência.

Figura 4: Resultados dos cenários aleatórios de convergência.

seu desempenho baseando-se principalmente na quantidade de iterações necessárias para convergir.

Apesar do Método CBPD possuir uma quantidade de iterações necessárias para convergir próximas do dobro de iterações quanto comparado com o Método de Newton, é válido levar em consideração o custo computacional envolvido no processamento de ambos os algoritmos. No método CBPD, para cada iteração possui uma complexidade $O(n)$ para ser calculada, considerando n a quantidade de parâmetros da função a ser otimizada. Já no Método de Newton, possui uma complexidade $O(n^3)$ decorrente da necessidade de inversão da matriz hessiana $n \times n$.

É válido seguir os estudos nessa pesquisa para analisar os critérios formais de convergência para o método caso existam, que podem estar relacionados com casos em que a matriz hessiana é diagonalizável e o valor dos elementos de sua matriz diagonal seja suficientemente próximos dos autovetores da matriz. Esse entendimento permite prever com mais assertividade os cenários em que é vantajoso a utilização do método.

Ademais é relevante analisar especificamente o caso de convergência para o ponto estacionário P-vdW, em que o método CBPD apresentou uma taxa de convergência consideravelmente superior ao do método de Newton, resultado esse não esperado, visto que o método de Newton tende a ser mais estável nos casos de convergência.

Referências

- [1] Felipe Fernandes Gomes da Silva Costa. Research Software Algorithm to Geometry Optimization, July 2023.
- [2] Jun Li, Richard Dawes, and Hua Guo. An ab initio based full-dimensional global potential energy surface for $\text{FH}_2\text{O}(\text{X}2\text{A}')$ and dynamics for the $\text{F} + \text{H}_2\text{O} \rightarrow \text{HF} + \text{HO}$ reaction. *The Journal of Chemical Physics*, 137(9):094304, 09 2012.
- [3] Jun Li, Bin Jiang, and Hua Guo. Permutation invariant polynomial neural network approach to fitting potential energy surfaces. ii. four-atom systems. *The Journal of Chemical Physics*, 13910:204103, 11 2013.
- [4] H. Bernhard Schlegel. Geometry optimization. *WIREs Computational Molecular Science*, 1(5):790–809, 2011.
- [5] J. Stewart. *CÁLCULO VOLUME 1 - TRADUÇÃO DA 6ª EDIÇÃO NORTE-AMERICANA*. Cengage Learning Edições Ltda., 2010.
- [6] J.V.C. Vargas and L.K. Araki. *Cálculo Numérico Aplicado*. MANOLE, 2016.
- [7] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.