

**UFRRJ**  
INSTITUTO DE CIÊNCIAS EXATAS  
MESTRADO ACADÊMICO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL  
**PPGMMC**

DISSERTAÇÃO

**MultiPatch: UM MÉTODO PARA TESSELAÇÃO DA SUPERFÍCIE  
EXCLUDENTE DO SOLVENTE**

FELIPE AUGUSTO DA SILVA



UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO  
INSTITUTO DE CIÊNCIAS EXATAS  
MESTRADO ACADÊMICO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL  
**PPGMMC**

**MultiPatch: UM MÉTODO PARA TESSELAÇÃO DA SUPERFÍCIE  
EXCLUDENTE DO SOLVENTE**

FELIPE AUGUSTO DA SILVA

Sob a Orientação do Professor  
Dulio Tadeu da Conceição Junior

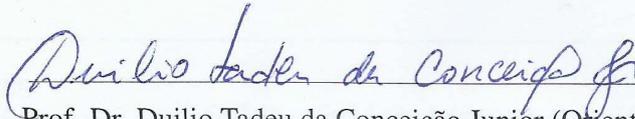
Dissertação submetida como requisito parcial  
para obtenção do grau de Mestre no Curso de  
Mestrado Acadêmico em Modelagem Matemá-  
tica e Computacional - PPGMMC.

Seropédica, RJ  
2016

**FELIPE AUGUSTO DA SILVA**

Dissertação submetida como requisito parcial para obtenção do grau de Mestre no Curso de Mestrado Acadêmico em Modelagem Matemática e Computacional - PPGMMC.

DISSERTAÇÃO APROVADA EM 29/9/2016



Prof. Dr. Duilio Tadeu da Conceição Junior (Orientador - UFRRJ)



Prof. Dr. Grigori Chapiro (Titular Externo - UFJF)



Prof. Dr. Márcio Soares Pereira (Titular Interno - UFRRJ)

005  
S586m  
T

Silva, Felipe Augusto da, 1987-  
Multipatch: um método para tesselação da  
superfície excludente do solvente / Felipe  
Augusto da Silva. - 2016.  
68 f.: il.

Orientador: Duilio Tadeu da Conceição  
Junior.

Dissertação (mestrado) - Universidade  
Federal Rural do Rio de Janeiro, Curso de  
Pós-Graduação em Modelagem Matemática e  
Computacional, 2016.

Bibliografia: f. 66-68.

1. Programas de computador - Teses. 2.  
Tesselações (Matemática) - Teses. 3.  
Moléculas - Modelos - Teses. 4. Solução  
(Química) - Modelos matemáticos - Teses.  
I. Conceição Junior, Duilio Tadeu da,  
1977- II. Universidade Federal Rural do  
Rio de Janeiro. Curso de Pós-Graduação  
em Modelagem Matemática e Computacional.  
III. Título.

*Dedico à minha mãe Marcia Maria da Silva (in memoriam), minha tia Ana Paula da Silva, meu avô Antônio Domingos da Silva e a minha avó Marli Dutra da Silva pela importância que dão a minha existência e a minha trajetória.*

## **AGRADECIMENTOS**

Agradeço a Deus pela experiência da vida e por me conceder uma família acolhedora.

Agradeço aos meus amigos próximos, aos companheiros de banda, Adilson Silva e Carlos Aguiar, aos amigos do Ciep 368 João Conceição Canuto, Colégio Estadual João Paulo II, Escola Municipal Coronel Moreira da Silva, Centro Educacional Congregacional. Muito obrigado pela paciência e compreensão quanto as minhas necessidades nesse período tão turbulento.

Agradeço aos meus orientadores Ph.D. Duílio Tadeu e Ph.D Clarissa Silva por todo o suporte intelectual e moral que me ofereceram.

Agradeço ao Ph.D Edgar Corrêa de Oliveira e ao Ph.D José Abdalla Helayel Neto pelo acolhimento e suporte durante o período de Iniciação Científica e pelo que continuam a fazer por mim.

Agradeço aos amigos do Programa de Mestrado em Modelagem Matemática e Computacional (PPGMMC): Bruna, Bruno, Carolina, Marlon, Maurício, Rúpila, Keitilane, Ulisses, Marcelli e Pablo. Fomos unidos até o fim e vencemos!

Agradeço à Janaina Gama, secretária do PPGMMC, por todo suporte e boa vontade.

## RESUMO

Entender o comportamento de moléculas em fase líquida é fundamental para o compreender diversos processos químicos e físicos, sendo assim referência para diversas áreas do conhecimento científico e tecnológico. Métodos usando o modelo de contínuo polarizável (PCM) tem sido propostos e revisados com o objetivo de melhorar a acuracia dos resultados e a aplicabilidade desse modelo. MultiPatch é um método analítico e eficiente para se obter uma tesselação da superfície molecular necessária para o computo de uma das componentes, de origem eletrostática, da energia de interação entre soluto e solvente  $\Delta G_s^*$ . O método foi testado comparando seus resultados para área da superfície e a energia  $\Delta G_s^*$  com o métodos internos do software GAMESS.

**Palavras-chave:** MultiPatch; tesselação de superfícies moleculares; modelo de contínuo polarizável

## ABSTRACT

To understand the behavior of molecules in liquid phase is fundamental for comprehend many chemical and physical process, became reference to several areas of scientific and tecnologic knowledge. Methods using polarizable continuum model (PCM) have been proposed and revised to achieve more acurated results and the aplicability of the method. MultiPatch is a analitic and efficient molecular surface tessellation method. This surface is necessary to calcule the electrostatic component of the interaction energy between solute and solvent  $\Delta G_s^*$ . The method was tested by comparing the results for the surface area and the energy  $\Delta G_s^*$  obtained using internal methods of the software GAMESS.

**Keywords: MultiPatch; molecular surfaces tessellation; polarizable continuum model**

---

## Lista de Figuras

---

|     |  |    |
|-----|--|----|
| 1.1 | Misturas homogêneas e heterogêneas. . . . .  | 16 |
| 1.2 | Dissociação do sal de cozinha em água. . . . .   | 17 |
| 1.3 | Betaína-30 diluída em vários solventes. . . . .  | 18 |
| 1.4 | Transferência da fase gasosa para o interior da solução. . . . .   | 18 |
| 1.5 | Modelos de solvente: a) Solvente como contínuo; b) Moléculas explícitas do solvente . . . . .  | 21 |
| 1.6 | Solvente polarizado. . . . .   | 22 |
| 3.1 | Nessa figura as bordas mais espessas representam a superfície da cavidade quando descrita por esferas de van der Waals (VDW), superfície acessível ao solvente (SAS), superfície excludente do solvente (SES). . . . .   | 40 |
| 3.2 | Processo de adição de esferas do GEPOL. (a) Para duas esferas próximas uma única esfera interceptando as esferas pais é gerada. (b) Para esferas mais distantes, primeiramente uma esfera que não intercepta as esferas 'pais' é gerada, então duas esferas de terceira geração são adicionadas. (c) Para qualquer par de esferas com afastamento maior, esferas menores bastante sobrepostas com as primitivas são geradas. As esferas pontilhadas representam as esferas de prova do solvente. . . . . | 41 |
| 3.3 | Processo de divisão equilátera . . . . .   | 42 |
| 4.1 | Divisão equilátera . . . . .   | 47 |
| 4.2 | Refinamento das esferas . . . . .  | 47 |
| 4.3 | Toro de revolução . . . . .  | 49 |
| 4.4 | Conectividade . . . . .  | 50 |
| 4.5 | Benzeno . . . . .  | 52 |
| 4.6 | Toro sobre uma molécula de CO . . . . .  | 52 |
| 4.7 | Patch de toro sobre uma molécula de CO . . . . .   | 53 |

|     |   |    |
|-----|---|----|
| 5.1 | Correlação da energia livre de solvatação experimental e numérica para as moléculas do grupo de controle: DMSO, Água, Metanol, Etanol. . . . .            | 57 |
| 5.2 | Comparação na energia livre de solvatação $\Delta G_s^*$ obtida a partir das cavidades geradas com o FIXPVA e o MultiPatch usando o método COSMO. . . . . | 61 |

---

## Lista de Tabelas

---

|     |  |    |
|-----|--|----|
| 5.1 | Grupo de Controle e energia de solvatação. . . . .             | 56 |
| 5.2 | Resultados utilizando o método DPCM não-iterativo . . . . .    | 58 |
| 5.3 | Resultados utilizando o método IEF-PCM não-iterativo . . . . . | 59 |
| 5.4 | Resultados utilizando o método COSMO não-iterativo . . . . .   | 60 |

|  |           |
|--|-----------|
| <b>Introdução</b>  | <b>12</b> |
| <b>1 SOLUÇÕES E SOLVATAÇÃO</b>                                   | <b>16</b> |
| 1.1 Misturas e soluções . . . . .                                | 16        |
| 1.2 Modelando a interação soluto-solvente . . . . .              | 18        |
| 1.2.1 Energia Livre de Solvatação $\Delta G_s^*$ . . . . .       | 18        |
| 1.2.2 Método de Campo de Reação . . . . .                        | 20        |
| <b>2 FORMULAÇÃO INTEGRAL E SOLUÇÕES NUMÉRICAS</b>                | <b>24</b> |
| 2.1 Formulação Integral da Equação de Poisson . . . . .          | 25        |
| 2.1.1 Equação de Laplace . . . . .                               | 25        |
| 2.1.2 Solução Fundamental . . . . .                              | 25        |
| 2.1.3 Equação de Poisson . . . . .                               | 27        |
| 2.1.4 Representação da solução por funções de Green . . . . .    | 27        |
| 2.1.5 Método de Equações Integrais . . . . .                     | 28        |
| 2.1.6 Modelos Contínuos por Formulações integrais . . . . .      | 29        |
| 2.1.7 Campo de Energia de Reação das Cargas Interiores . . . . . | 31        |
| 2.1.8 Métodos de Discretização . . . . .                         | 33        |
| 2.1.9 Aspectos Numéricos . . . . .                               | 37        |
| <b>3 DISCRETIZAÇÃO DA SES</b>                                    | <b>39</b> |
| 3.1 GEPOL . . . . .  | 40        |
| 3.1.1 Partição das interseções entre esferas . . . . .           | 42        |
| 3.1.2 Dificuldades . . . . .                                     | 42        |
| 3.2 DefPol . . . . .   | 43        |
| 3.3 Conclusões . . . . .   | 44        |

---

|          |  |           |
|----------|--|-----------|
| <b>4</b> | <b>MultiPatch</b>                        | <b>46</b> |
| 4.1      | Tesselação dos objetos básicos . . . . . | 46        |
| 4.1.1    | Esfera . . . . .                         | 46        |
| 4.1.2    | Toro . . . . .                           | 48        |
| 4.2      | Patches . . . . .                        | 51        |
| 4.2.1    | Esferas Covexas . . . . .                | 51        |
| 4.2.2    | Patch de Toro . . . . .                  | 51        |
| 4.2.3    | Patch de Esfera concava . . . . .        | 53        |
| 4.2.4    | Mecanismos de remoção . . . . .          | 54        |
| 4.2.5    | Implementação . . . . .                  | 54        |
| <b>5</b> | <b>Resultados Numéricos</b>              | <b>55</b> |
| 5.1      | GAMESS . . . . .                         | 55        |
| 5.2      | Análise de dados . . . . .               | 56        |
| 5.2.1    | Testes com GEPOL . . . . .               | 56        |
| 5.2.2    | Testes com MultiPatch . . . . .          | 56        |
| <b>6</b> | <b>Conclusões</b>                        | <b>62</b> |
| <b>A</b> | <b>Entrada e saída do GAMESS</b>         | <b>63</b> |

---

---

## INTRODUÇÃO

---

Com os avanços tecnológicos relacionados a computação nas duas últimas décadas diversos modelos teóricos associados à soluções líquidas puderam ser implementados, testados, revistos e aperfeiçoados. O uso de soluções numéricas possibilitou que sistemas ou processos mais complexos, por vezes com tratamento analítico inviável ou impossível, fossem explorados de maneira cada vez mais acurada e eficiente.

Em uma solução, o composto químico mais abundante é chamado de solvente enquanto os demais compostos são chamados de solutos. Solvatação é a interação estabilizadora entre um íon ou molécula de soluto e as moléculas do solvente formando um arranjo específico [24].

Dentre os diversos modelos propostos para descrever a influência do solvente sobre o soluto [11, 17, 26], os modelos de contínuo apresentam o paradigma ideal para a descrição dessa influência no nível quântico [22]. Modelos de contínuo quanto-mecânicos básicos são tais que respeitam as seguintes características:

1. O soluto é descrito no nível quanto-mecânico.
2. Apenas a interação eletrostática entre soluto e solvente é considerada. Outras formas de interação, quando presentes, são descritas por leis empíricas.
3. A solução é muito diluída, ou seja, apenas uma molécula do soluto encontra-se solvatada pelo solvente.
4. O solvente é isotrópico e em equilíbrio para uma dada temperatura.
5. Apenas a configuração eletrônica de menor energia do soluto é considerada (*ground state*).
6. Nenhum efeito dinâmico sobre a solução é considerado.

---

Para a alocação da molécula do soluto define-se uma cavidade vazia no interior do meio contínuo dielétrico que representa as moléculas do solvente. A forma e o tamanho dessa cavidade são definidas de formas diferentes para cada versão de modelo contínuo, entretanto deve conter praticamente toda distribuição de cargas do soluto.

Uma cavidade de particular interesse é a superfície excludente do solvente (SES) definida como a região do espaço onde uma prova esférica do solvente não pode entrar sem que intercepte regiões esféricas com centros e raios relativos aos átomos que compõem o soluto ??.

A distribuição de cargas do soluto  $\rho_M$  induz uma polarização no solvente que por sua vez induz uma polarização na distribuição de cargas do soluto.

Para se determinar a configuração de equilíbrio da solução é necessário adicionar a influência do potencial eletrostático gerado a partir dessas polarizações. Esse potencial é referenciado como potencial de reação  $V_R$  [11, 22], definido como a diferença entre o potencial eletrostático gerado por  $\rho_M$  em fase gasosa (ou no vácuo) e quando em solução. Faz-se necessário portanto a solução de um problema eletrostático básico descrito pela equação de Poisson [23].

Uma das técnicas para se resolver o problema eletrostático é referenciada como método da superfície de cargas aparente (ASC). Nesse método uma distribuição de cargas superficial  $\sigma$  é estabelecida na superfície  $\Gamma$  da cavidade  $C$ . O potencial gerado por essa distribuição de cargas corresponde exatamente ao potencial de reação  $V_R$  descrito acima. Modelos de meios contínuos dielétricos polarizáveis como DPCM e IEFPCM [22] e modelos de meios condutores como o COSMO e CPCM [19] se utilizam do paradigma ASC.

Para se determinar efetivamente a distribuição  $\sigma$  faz-se necessário o uso de métodos de elementos de superfície. Métodos para se obter a discretização da cavidade molecular como GEPOL [25], DEFPOL [7], FIXPVA, GEPOL-AS, GEPOL-RT E GEPOL-GB [30] foram propostos, implementados e testados, revelando os aspectos positivos e negativos de cada técnica.

Em geral, assume-se que a distribuição  $\sigma$  é composta pela soma de funções constantes em cada elemento de superfície, também chamado de tessera [22].

Atualmente o método GEPOL [25] é o mais amplamente utilizado para se obter a tesselação sobre a SES. Entretanto, ainda são relatados alguns problemas de instabilidade numérica [7], principalmente em relação a obtenção das derivadas, necessárias para otimização da geometria molecular. A tesselação obtida com o FIXPVA [30] resolve a questão das derivadas suavizando as áreas das tesseras em função da sua distância entre as outras esferas centradas nos átomos do soluto (*area scaling*), mas a superfície gerada não é necessariamente a SES.

A partir da análise dos métodos citados acima obtivemos o conjunto de características a serem respeitadas no intuito de se obter uma nova metodologia.

No método MultiPatch, propomos aproximar a SES utilizando objetos geométricos que de fato definem a SES da forma como é definida, de modo a minimizar os erros numéricos devidos a aproximação da superfície.

Após o desenvolvimento e implementação do algoritmo, foram realizados experimen-

---

tos numéricos onde comparamos os resultados para a energia livre de solvatação  $\Delta G_s^*$  obtidos experimentalmente [21] com os obtidos pelo software GAMESS [27], sendo a SES obtida pelo GEPOL/FIXPVA [30], próprio do GAMESS, e pelo MultiPatch.

Mantendo o fluxo de informações como no exposto acima, o texto foi organizado da seguinte forma:

O capítulo 1 contém uma motivação geral para esse trabalho. É definido o fenômeno físico-químico da solvatação, sua importância e alguns exemplos do efeito desse fenômeno.

O capítulo 2 contém a argumentação que permite a obtenção de uma formulação integral para o cálculo da energia livre de solvatação e os métodos de discretização usuais.

No capítulo 3 são definidos os tipos de cavidade molecular, em especial a necessária para o computo da interação eletrostática entre soluto e solvente, a SES. Os métodos GEPOL [25] e DefPol [7] serão apresentados assim como suas dificuldades técnicas.

No capítulo 4 é apresentada a metodologia MultiPatch, definindo-se a tesselação dos objetos básicos e os mecanismos de remoção para tesseras em excesso.

Os capítulos 5 e 6 contém os resultados dos experimentos numéricos e as conclusões obtidas a partir da análise dos dados experimentais, respectivamente.

# CAPÍTULO 1

## SOLUÇÕES E SOLVATAÇÃO

### 1.1 Misturas e soluções

Misturas entre substâncias são muito comuns tanto no dia a dia quanto em procedimentos técnicos e experimentais específicos. São usualmente classificadas como homogêneas quando apresentam apenas uma fase, como a mistura entre água e álcool, ou heterogênea quando apresentam mais de uma fase, como no caso da mistura de água com óleo de cozinha (Figura 1.1). Uma fase é cada porção que apresenta uniformidade quanto a composição química e estado físico.

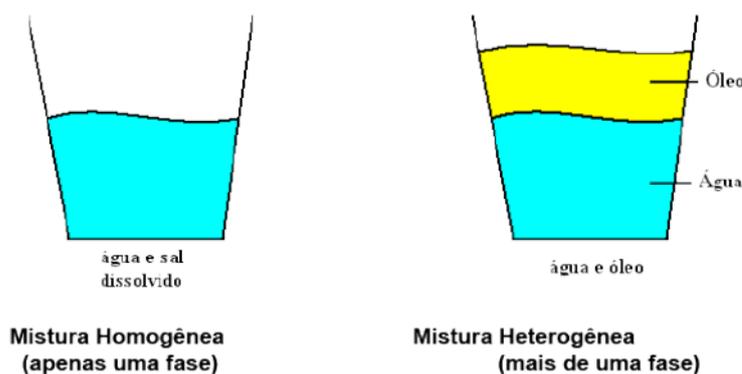


Figura 1.1: Misturas homogêneas e heterogêneas.

Em geral, chamamos de solvente a substância em maior quantidade e de soluto as de menor quantidade. Algumas características físicas da água como a constante dielétrica, calor de vaporização, calor específico e calor latente de fusão fazem com que ela seja considerada um solvente universal.

Soluções são misturas homogêneas cujo tamanho das partículas que compõem o sol-

vente e os solutos é inferior a 1nm e podem ocorrer em todos os estados de agregação.

Ao se analisar a dissolução de um soluto leva-se em consideração duas etapas: a dissociação e a solvatação. A primeira diz respeito a separação de moléculas ou íons do soluto e a segunda tem que ver com a manutenção dessa separação. Em outras palavras, solvatação é o processo em que um íon ou molécula de soluto é envolvido pelas moléculas do solvente arranjadas de uma maneira específica. No caso particular onde o solvente é água, é usual chamar esse fenômeno como hidratação [24].

Um exemplo simples de um sistema solvatado é a dissociação iônica do NaCl quando dissolvido em água (Figura 1.2).

Do ponto de vista molecular-microscópico, solventes podem “quebrar” cadeias de cristais em reagentes sólidos, dissolver reagentes gasosos ou líquidos e exercer uma considerável influência sobre a quantidades de reagentes e produtos de uma reação no equilíbrio químico, alterar a energia livre de ativação assim como os produtos de uma reação, devido à solvatação dos reagentes/estados de transição possibilita a transferência de prótons para reações sujeitas a catálise ácido-base, entre outras. Temos como exemplo a influência sobre as taxas de heterólise<sup>1</sup> unimolecular de 2-cloro-2 metilpropano observada em água e em benzeno, que aumenta por um fator de aproximadamente  $10^{11}$  quando o benzeno não-polar é substituído por água (polar) [26, 11].

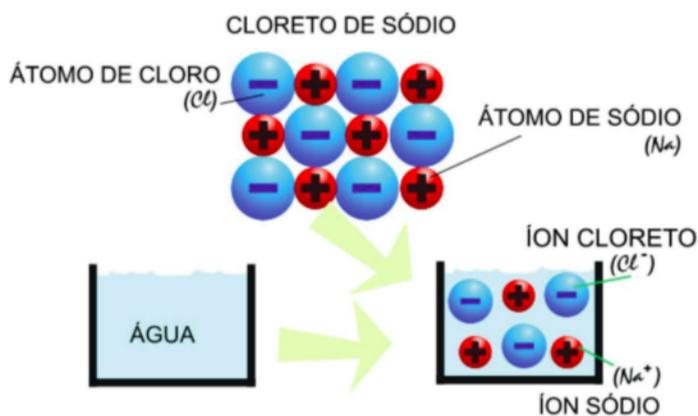


Figura 1.2: Dissociação do sal de cozinha em água.

A seleção adequada do solvente ou mistura de solventes não é importante apenas para processos químicos mas também em processos físicos, tais como a recristalização, todos os processos de extração, particionamento, separações cromatográficas, reações catalíticas de transferência de fase, etc. De particular interesse, neste contexto, é a influência de solventes em todos os tipos de processos de absorção de luz, por exemplo, em UV/visível, infravermelha e os espectros de ressonância magnética nuclear (RMN). A dependência do solvente na absorção UV/visível é referenciada como *solvatocromismo* (Figura 1.3). Corantes solvatocrômicos podem ser utilizados como indicadores empíricos de polaridade do solvente.

<sup>1</sup>Quebra da ligação química gerando espécies iônicas.

Sobre solventes líquidos cabe destacar que macroscopicamente são considerados como meios ideais para transportar calor e massa em reações exotérmicas e endotérmicas. Outro fato é que meios líquidos podem dissolver os reagentes e ser facilmente separados do produto da reação posteriormente, portanto o solvente pode ter uma influência decisiva no resultado de uma reação química sob estudo.



Figura 1.3: Betaína-30 diluída em vários solventes.

## 1.2 Modelando a interação soluto-solvente

### 1.2.1 Energia Livre de Solvatação $\Delta G_s^*$

Um parâmetro que descreve bem a interação soluto-solvente e está associado ao potencial químico do soluto é chamado de energia livre de solvatação  $\Delta G_s^*$  [1], representando a variação da energia livre de Gibbs quando um íon ou molécula  $s$  é transferido do vácuo (fase gasosa) para o interior do solvente em uma dada posição fixa  $\mathbf{R}_0$ . Essa propriedade termodinâmica pode ser interpretada como o trabalho  $W(s|sol)$  necessário para se realizar essa transferência.

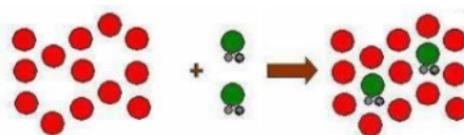


Figura 1.4: Transferência da fase gasosa para o interior da solução.

### Relação Fundamental da Termodinâmica para soluções

Uma descrição adequada do estado termodinâmico de uma mistura deve conter grandezas relacionadas a quantidade de cada componente, como sua massa  $m$  ou número de mols  $N$ , assim como um parâmetro que indique a variação da energia do sistema quanto ao incremento de cada componente no sistema. Esse parâmetro é chamado de potencial químico  $\mu$ .

A relação fundamental da Termodinâmica [5] correspondente é dada pela energia interna  $U(S, V, N_1, \dots, N_n)$  cujo diferencial é:

$$dU = TdS - pdV + \sum_{i=1}^n \mu_i dN_i \quad (1.1)$$

onde  $T$  é a temperatura absoluta,  $S$  é a entropia,  $p$  é a pressão e  $N_i$  e  $\mu_i$  representam, respectivamente, o número de mols e o potencial químico do  $i$ -ésimo componente da solução.

Usando uma transformação de Legendre apropriada [5], outras representações podem ser obtidas. Entre elas temos a representação pela energia livre de Gibbs  $G(T, p, N_1, \dots, N_n)$  cujo diferencial é dado por:

$$dG = -SdT + Vdp + \sum_{i=1}^n \mu_i dN_i \quad (1.2)$$

a partir da qual se obtém

$$S = \frac{\partial G}{\partial T}; \quad V = \frac{\partial G}{\partial p}; \quad \mu_i = \frac{\partial G}{\partial N_i} \quad (1.3)$$

Entretanto, os argumentos da Termodinâmica não fornecem a dependência funcional do potencial químico sobre a temperatura, pressão ou composição do sistema. Em princípio, esta é obtida a partir da Mecânica Estatística.

Consideramos, por simplicidade, um sistema com apenas dois componentes e que o incremento de uma única molécula em um sistema macroscópico é infinitesimal. Dado o caráter extensivo<sup>2</sup> da energia de Gibbs [1] é possível representar o potencial químico  $\mu_A$  do componente A após a adição de uma molécula<sup>3</sup> como:

$$\mu_A = G(T, p, N_A + 1, N_B) - G(T, p, N_A, N_B) \quad (1.4)$$

Uma quantidade útil para auxiliar na interpretação dos fatores que influenciam no potencial químico é o *potencial pseudo-químico*  $\mu_A^*$ , que representa a adição de uma molécula em uma posição fixa  $\mathbf{R}_0$ , dado por

$$\mu_A^* = G(T, p, N_A + 1, N_B; \mathbf{R}_0) - G(T, p, N_A, N_B) \quad (1.5)$$

A partir de argumentos da Mecânica Estatística [1] obtém-se uma importante relação envolvendo as quantidades descritas acima dada por:

$$\mu_A = \mu_A^* + kT \ln(\rho_A \Lambda_A^3) \quad (1.6)$$

De maneira geral, o primeiro termo do lado direito dessa expressão é interpretado como a energia de acoplamento  $W(s|sol)$  da molécula  $s$  de soluto no solvente somada a outros

<sup>2</sup>Parâmetros que, em um sistema composto, tem valores iguais a soma dos valores em cada subsistema são chamados parâmetros extensivos.

<sup>3</sup>Um mol de uma substância contém aproximadamente  $6,02 \cdot 10^{23}$  moléculas de uma substância, portanto a adição de uma única molécula pode ser pensada como um incremento praticamente infinitesimal.

efeitos devido as moléculas do solvente que a envolve. O segundo termo da direita é chamado de energia de liberação, uma vez que remover o vínculo sobre a posição  $\mathbf{R}_0$  de  $s$  torna possível uma mudança na energia livre na forma de energia cinética, onde  $\rho_s$  é a densidade numérica do componente  $s$  e  $\Lambda_A^3$  é referenciado como a função de partição de momento linear [1].

A energia necessária para se transportar  $s$  de uma posição fixa na fase gasosa ideal ( $gi$ ) para uma posição fixa em solução ( $sol$ ) é chamada de energia livre de solvatação de Gibbs e é definida como:

$$\Delta G_s^* = \mu_s^{*sol} - \mu_s^{*gi} \quad (1.7)$$

Quando não há interação entre as partículas, como no caso dos gases ideais, o potencial pseudo-químico pode ser reduzido a:

$$\mu_s^{*gi} = -kT \ln(q_n) \quad (1.8)$$

onde  $q_n$  é a função de partição interna (rotacional, vibracional, eletrônica e nuclear)[5].

Outro caso particular simples se dá quando a partícula solvatada não tem graus de liberdade internos ou quando esses não são afetados pelo solvente que a envolve. Em ambos os casos podemos escrever

$$\mu_s^{*sol} = W(s|sol) - kT \ln(q_n) \quad (1.9)$$

Pelas Equações (1.7), (1.8) e (1.9), temos que a energia livre de solvatação resulta exatamente na energia de acoplamento, ou seja,

$$\Delta G_s^* = W(s|sol) \quad (1.10)$$

## 1.2.2 Método de Campo de Reação

De maneira geral, a envoltória de moléculas de solvente ao redor soluto é idealizada através de moléculas explícitas ou como um meio contínuo homogêneo que é polarizado pela presença do soluto [11]. No Método de Campo de Reação [6], que é amplamente utilizado no estudo dos efeitos do solvente, adota-se o segundo paradigma. A região ocupada pelo soluto é chamada de cavidade molecular  $C$  e sua superfície é representada por  $\Gamma = \partial C$ <sup>4</sup>.

A interação soluto-solvente é descrita nesse método em termos das componentes de cavitação<sup>5</sup>  $G_{cav}^*$ , eletrostática  $G_{el}^*$  e repulsiva-dispersiva<sup>6</sup>  $G_{r-p}^*$ , além da contribuição devido aos movimentos térmicos do sistema de referência [11].

<sup>4</sup>Dada uma região aberta  $R$ , é usual representar o contorno de  $R$  por  $\partial R$

<sup>5</sup>Cavitação está relacionada a formação no solvente puro de uma cavidade vazia com o formato e tamanho apropriado para acomodar o soluto.

<sup>6</sup>A distâncias próximas ao raio nuclear, mesmo partículas com cargas elétricas opostas interagem de maneira a se repelirem, oscilando entre uma posição de equilíbrio instável. A componente dispersiva resulta da formação temporária de dipolos dada a posição de dois átomos adjacentes.

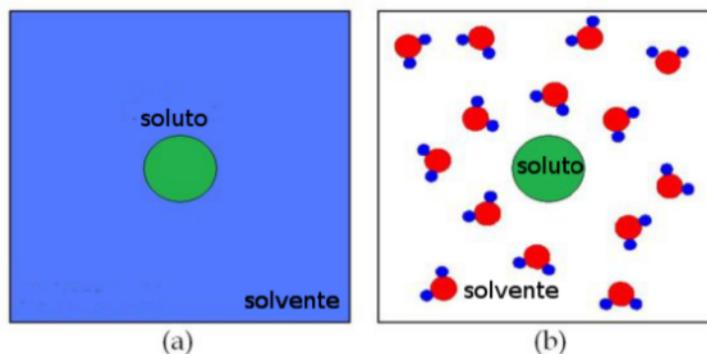


Figura 1.5: Modelos de solvente: a) Solvente como contínuo; b) Moléculas explícitas do solvente

$$\Delta G_s^* = \Delta G_{cav}^* + \Delta G_{el}^* + \Delta G_{r-p}^* + \Delta G_{tm}^* \quad (1.11)$$

A ordem das contribuições indicadas na equação (1.11) corresponde a melhor sequência na qual o processo de acomodação do soluto pode ser desenvolvido [11].

Na versão quanto-mecânica do método, a molécula é estudada *ab initio* e a interação soluto-solvente é obtida através de um potencial de reação  $\Phi_R$ , agindo como uma perturbação na Hamiltoniana do soluto [6]

$$\begin{aligned} H^0 \Psi^0 &= E^0 \Psi^0 && \text{no vácuo,} \\ [H^0 + \Phi_R] \Psi &= E \Psi && \text{em solução.} \end{aligned} \quad (1.12)$$

onde  $H^0$  é a Hamiltoniana do soluto no vácuo,  $\Psi^0$  e  $\Psi$  são as funções de onda do soluto no vácuo e em solução, respectivamente.

A componente eletrostática é descrita em termos de um conjunto de cargas pontuais induzidas  $\{q_i\}$  posicionadas no centro de pequenos elementos de área, *tesseras*, que cobrem a superfície  $\Gamma$  da cavidade onde o soluto está inserido. No Capítulo 4 serão apresentadas algumas definições para essa superfície, sendo a Superfície Excludente do Solvente (SES) a motivação para esse trabalho. As demais interações, quando incluídas, são determinadas através de modelos empíricos ou por outras aproximações e não serão discutidas nesse texto [26].

Determinamos o trabalho  $W$  necessário para se alocar uma distribuição de cargas  $\rho$  através do seguinte procedimento: Consideramos que a alocação é formada trazendo elementos infinitesimais de carga do infinito até suas posições finais. Em um certo instante a distribuição de cargas assume uma configuração  $\lambda\rho$ , onde  $\lambda$  é um parâmetro que varia de 0 a 1. O potencial eletrostático  $\Phi(\mathbf{r})$ , que é uma função linear em relação a distribuição de cargas [23], assume nesse mesmo instante um valor  $\lambda\Phi$ , onde  $\Phi$  é o valor final desse potencial. Um incremento  $\delta\lambda$  é equivalente a trazer  $\rho\delta\lambda dv$  elementos de carga do infinito para dentro do volume  $V$ , fazendo a carga total aumentar de  $\int_{\mathbb{R}^3} \lambda\rho dv$  até  $\int_{\mathbb{R}^3} (\lambda+d\lambda)\rho dv$ . Para que um elemento de carga  $\rho\delta\lambda dv$  seja colocado numa região de potencial  $\lambda\Phi$ , é necessário uma quantidade de trabalho  $(\rho\delta\lambda dv)\lambda\Phi$ .

Assim, o trabalho necessário para aumentar a distribuição de carga em  $\rho\delta\lambda$  é:

$$\delta W = \int_{\mathbb{R}^3} \lambda d\lambda \rho \Phi dv \quad (1.13)$$

Esse procedimento é válido para todos os elementos de carga. Assim, tomando a integral de  $\delta W$  entre  $\lambda = 0$  e  $\lambda = 1$ , teremos

$$W = \frac{1}{2} \int_{\mathbb{R}^3} \rho \Phi dv \quad (1.14)$$

Caso essa acomodação ocorra no interior de um meio contínuo dielétrico, como assumido para o solvente, o campo elétrico  $\mathbf{E}$  gerado por essa distribuição de cargas induz a polarização desse meio com distribuição de carga  $\rho'$ . Um campo vetorial de polarização  $\mathbf{P}$  simula o efeito dos dipolos explícitos das moléculas do solvente, obtido como uma aproximação de resposta linear ao campo elétrico  $\mathbf{E}$  [23, 8], ou seja,

$$\mathbf{P} = \chi \mathbf{E} \quad (1.15)$$

onde  $\chi$  representa a suscetibilidade elétrica do meio.

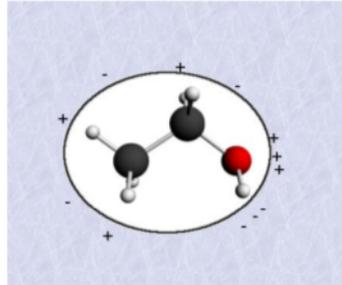


Figura 1.6: Solvente polarizado.

De maneira similar à equação (1.14), a energia de interação entre as distribuições de carga  $\rho$  e  $\rho'$  é dada por

$$E_R(\rho, \rho') = \int_{\mathbb{R}^3} \rho'(\mathbf{r}) \Phi_R(\mathbf{r}) d\mathbf{r} \quad (1.16)$$

onde  $\Phi_R$  representa a diferença entre o potencial eletrostático gerado pela distribuição  $\rho'$  no interior do dielétrico e no vácuo.

Sendo  $\mathbf{E}$  é um campo conservativo [23] existe  $\Phi$  tal que:

$$\mathbf{E} = -\nabla\Phi \quad (1.17)$$

Aplicando-se a lei de Gauss da eletrostática sobre uma superfície fechada dentro do dielétrico, considerando  $\rho(\mathbf{r})$  a densidade de carga externa a superfície de Gauss e o teorema da divergência [28], é possível estabelecer que:

$$\nabla \cdot [\epsilon(\mathbf{r}) \nabla \Phi(\mathbf{r})] = -\rho(\mathbf{r}). \quad (1.18)$$

Essa expressão é conhecida como equação de Poisson e  $\epsilon$  é um tensor que representa a permissividade elétrica do meio, sendo  $\epsilon(\mathbf{r}) = 1$  no vácuo.

Com os elementos apresentados até aqui já seria possível obtermos soluções aproximadas para  $\Phi = \Phi_R + \Phi_{\rho'}$  por algum método numérico, sendo  $\Phi_{\rho'}$  o potencial devido a distribuição  $\rho'$  no vácuo. Com a distribuição  $\rho'$  dada,  $\Phi_{\rho'}$  pode ser calculada e usada para aproximar  $\rho$ , donde calcula-se  $\Phi$ . Iterativamente aproximamos  $\Phi_R = \Phi - \Phi_{\rho'}$  e conseqüentemente  $\Delta G_{el}^*$  por (1.16).

Existe, porém, outra abordagem para o cálculo do potencial  $\Phi_{\rho'}$  mais adequada para cálculos quanto-mecânicos. Nessa abordagem a equação (1.16) é reescrita com uma formulação integral cuja solução depende não mais da distribuição de cargas volumétrica  $\rho'$  e sim da distribuição de cargas superficial  $\sigma$ .

## CAPÍTULO 2

---

# FORMULAÇÃO INTEGRAL E SOLUÇÕES NUMÉRICAS

---

No capítulo anterior vimos de forma esquematizada como determinar o campo de energia potencial  $E_R$  e a partir dele calcular a energia livre de solvatação  $\Delta G_s^*$

Nosso objetivo agora é apresentar o conjunto de ferramentas matemáticas necessário para reformular a equação de Poisson através de equações integrais, donde derivam os métodos da família PCM (DPCM, IEF-PCM, COSMO...[22]), e os métodos de discretização usuais para esse tipo de formulação.

O primeiro passo é reescrever o potencial de reação

$$E_R(\rho, \rho') = \int_{\mathbb{R}^3} \rho'(\mathbf{r}) \Phi_R d\mathbf{r} \quad (2.1)$$

como uma integral sobre a superfície  $\Gamma = \partial C$ , onde  $C$  é a cavidade ocupada pelo soluto, assumindo a forma

$$E_R(\rho, \rho') = \int_{\Gamma} \sigma \Phi'_M d\Gamma \quad (2.2)$$

onde  $\Phi'_M$  é o potencial gerado pela distribuição de cargas  $\rho'$  no vácuo, ou seja,

$$\Phi'_M = \int_{\mathbb{R}^3} \frac{\rho'(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (2.3)$$

e a superfície de carga  $\sigma$  é uma solução de uma equação integral sobre  $\Gamma$ , calculada por:

$$\int_{\Gamma} k_A(\mathbf{s}, \mathbf{s}') \sigma(\mathbf{s}') ds' = b_\rho(\mathbf{s}), \forall \mathbf{s} \in \Gamma \quad (2.4)$$

onde  $k_A$  é o núcleo de Green de um operador integral  $A$ , determinado para cada método da família PCM, e o lado esquerdo  $b_\rho$  depende linearmente da distribuição de cargas  $\rho$ .

Essa transformação simplifica demasiadamente o problema eletrostático, pois a fonte do potencial de reação é reduzida a distribuição superficial de cargas  $\sigma$  sobre a superfície  $\Gamma$  apenas.

Apesar da simplificação no formalismo, a integração da equação (2.4) sobre uma superfície de forma complexa é um desafio computacional. As soluções são geralmente baseadas na discretização da integral em um número finito de elementos. A superfície da cavidade  $\Gamma$  é aproximada por um conjunto finito de elementos (chamados de tesseras) pequenos o suficiente para que  $\sigma$  possa ser considerado praticamente constante sobre cada tessera. Com  $\sigma$  completamente definido ponto a ponto, podemos definir um conjunto de cargas pontuais  $\{q_k\}$ , tal que  $q_i = \sigma_i A_i$  sobre cada tessera  $t_i$ .

Assumimos que o soluto estará completamente contido no interior da cavidade  $C$ . Caso contrário essa representação não é formalmente válida [22], como é no caso, por exemplo, de soluções iônicas.

## 2.1 Formulação Integral da Equação de Poisson

### 2.1.1 Equação de Laplace

A equação de Laplace é definida por

$$\nabla^2 u = 0 \quad \text{em } \Omega \in \mathbb{R}^n, \quad (2.5)$$

onde  $\nabla^2 = \partial^2/\partial x_1^2 + \dots + \partial^2/\partial x_n^2$  é o operador de Laplace. As soluções da equação de Laplace são chamadas de funções harmônicas, cujo estudo é o objetivo da teoria de potencial. Tais soluções são de grande importância em diversas áreas da ciência, notadamente em eletromagnetismo.

O conjunto  $\Omega \subset \mathbb{R}^n$  é um conjunto aberto conexo por caminhos denominado *domínio*. Um conjunto *conexo* é um conjunto tal que para quaisquer dois pontos no conjunto existe uma curva contínua contida no conjunto e ligando os dois pontos.

Observe que a condição da derivada segunda se aplica apenas sobre  $\Omega$  e não sobre seu contorno  $\Gamma = \partial\Omega$ . Para que valores atribuídos sobre o contorno façam sentido, devemos assumir que  $u$  seja contínua em  $\bar{\Omega} = \Omega \cup \Gamma$ , o que conduz à seguinte definição:

**Definição 2.1.1.** *Uma função  $u$  é dita harmônica em  $\Omega$  se  $u \in C^2(\Omega) \cup C^0(\bar{\Omega})$  e satisfaz a equação (2.5).*

### 2.1.2 Solução Fundamental

Definimos a função de singularidade  $s$  como:

$$s(\mathbf{x}, \mathbf{y}) := \begin{cases} -\frac{1}{\omega_2} \ln |\mathbf{x} - \mathbf{y}| & \text{para } n = 2 \\ \frac{|\mathbf{x} - \mathbf{y}|^{2-n}}{(n-2)\omega_n} & \text{para } n > 2 \end{cases} \quad (2.6)$$

para  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , onde

$$\omega_n = n \frac{\pi^{\frac{n}{2}}}{\Gamma(1 + \frac{n}{2})} \quad (2.7)$$

representa a área da superfície da esfera unitária  $n$ -dimensional [3]. Para  $n = 3$  temos  $\omega_3 = 4\pi$ . A norma nesse caso é a norma Euclidiana

$$|x| = \left( \sum_{i=1}^n x_i^2 \right)^{1/2}. \quad (2.8)$$

**Lema 2.1.2.** Dado  $y \in \mathbb{R}^n$  fixado, a função  $s(x, y)$  é uma solução da equação de Laplace em  $\mathbb{R}^n \setminus \{y\}$ .

**Demonstração.** Vide [32].

**Definição 2.1.3.** A derivada normal de  $u$  em  $\mathbf{x} \in \Gamma$  é definida como  $\frac{\partial u(\mathbf{x})}{\partial n} = \langle \mathbf{n}(\mathbf{x}), \nabla u(\mathbf{x}) \rangle$  onde  $\nabla u$  é o gradiente de  $u$  e  $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i$  é o produto escalar em  $\mathbb{R}^n$

Seja  $u \in C^1(\bar{\Omega})$  e  $v \in C^2(\bar{\Omega})$ , onde o domínio  $\Omega$  satisfaz determinadas condições [18], então é válida a primeira fórmula de Green:

$$\int_{\Omega} u \nabla^2 v d\mathbf{x} = - \int_{\Omega} \langle \nabla u, \nabla v \rangle + \int_{\partial\Omega} u \frac{\partial v}{\partial n} d\Gamma \quad (2.9)$$

Um domínio tal que a Equação (2.9) vale é chamado de *domínio normal*.

Funções  $u, v \in C^2(\bar{\Omega})$  em um domínio normal  $\Omega$  satisfazem a segunda fórmula de Green

$$\int_{\Omega} u \nabla^2 v d\mathbf{x} = \int_{\Omega} v \nabla^2 u + \int_{\partial\Omega} \left[ u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \right] d\Gamma \quad (2.10)$$

**Teorema 2.1.4.** Seja  $\Omega$  um domínio normal. Se  $u \in C^2(\bar{\Omega})$  é uma função harmônica em  $\Omega$ , então

$$\mathbf{u}(\mathbf{y}) = \int_{\partial\Omega} \left[ s(\mathbf{x}, \mathbf{y}) \frac{\partial \mathbf{u}(\mathbf{x})}{\partial n} - u(\mathbf{x}) \frac{\partial s(\mathbf{x}, \mathbf{y})}{\partial n_x} \right] d\Gamma_x \quad \forall \mathbf{y} \in \Omega. \quad (2.11)$$

**Demonstração.** Vide [32].

Se para um  $y \in \Omega$  fixado a função  $\phi(\cdot, \mathbf{y})$  é harmônica em  $\Omega$  e pertence a  $C^2(\bar{\Omega})$ , então a função

$$\gamma(x, y) = s(\mathbf{x}, \mathbf{y}) + \phi(\mathbf{x}, \mathbf{y}) \quad (2.12)$$

também satisfaz o Lema 2.1.2. Tal função é chamada solução fundamental da equação de Laplace em  $\mathbf{x} \in \Omega$ . Usualmente considera-se como solução fundamental a própria função  $s(\mathbf{x}, \mathbf{y})$ , uma vez que  $\phi(\mathbf{x}, \mathbf{y})$  identicamente nula é uma função harmônica.

**Corolário.** *Sob as hipóteses do Teorema 2.1.4, para toda solução fundamental  $\gamma$  em  $\Omega$  é válido*

$$\mathbf{u}(\mathbf{y}) = \int_{\partial\Omega} \left[ \gamma(\mathbf{x}, \mathbf{y}) \frac{\partial \mathbf{u}(\mathbf{x})}{\partial n} - u(\mathbf{x}) \frac{\partial \gamma(\mathbf{x}, \mathbf{y})}{\partial n_x} \right] d\Gamma_x, \quad \forall \mathbf{y} \in \Omega. \quad (2.13)$$

**Demonstração.** *Vide [32].*

### 2.1.3 Equação de Poisson

A equação de Poisson é definida como

$$\nabla^2 u = f \text{ em } \Omega \in \mathbb{R}^n \quad (2.14)$$

para um dado  $f \in C^0(\Omega)$ . Fisicamente, o termo  $f$  é um termo de fonte, como por exemplo, uma densidade de carga no caso de um potencial elétrico  $u$ . É importante notar que a solução da equação de Poisson (2.14) não é única. Para tal, é necessário especificar mais uma condição sobre a solução  $u$ . Tal condição é uma condição de fronteira, que para o momento consideramos a condição de Dirichlet que é:

$$u = \varphi \text{ sobre } \Gamma. \quad (2.15)$$

onde  $\varphi$  é uma função conhecida, o qual especifica o valor de  $u$  sobre a fronteira do domínio. Discutiremos em outro momento outras condições de fronteira.

Uma função  $u$  em  $C^2(\Omega) \cap C^0(\bar{\Omega})$  que satisfaz a equação (2.14) e a condição de fronteira é chamada solução clássica.

### 2.1.4 Representação da solução por funções de Green

Dada uma equação de Poisson (2.14) com condição de Dirichlet prescrita na fronteira do domínio, é possível representar sua solução utilizando integrais e uma solução fundamental  $\gamma$ , como apresentado no seguinte Lema:

**Lema 2.1.5.** *Vamos supor que  $u \in C^2(\Omega)$  é solução da equação de Poisson com condição de Dirichlet prescrita em (2.15) e  $\Omega$  é um domínio normal. Se  $\gamma$  é uma solução fundamental da equação de Laplace em  $\Omega$ , então  $u$  pode ser representada por*

$$u(x) = - \int_{\Omega} \gamma(\xi, x) f(\xi) d\xi + \int_{\Gamma} \left[ \gamma(\xi, x) \frac{\partial u(\xi)}{\partial n} - u(\xi) \frac{\partial \gamma(\xi, x)}{\partial n_\xi} \right] d\Gamma_\xi \quad (2.16)$$

para todo  $\gamma$  em (2.12).

**Demonstração.** Vide [32].

**Definição 2.1.6.** Uma solução fundamental  $G$  é chamada de função de Green (de primeiro tipo) se  $G(\xi, x) = 0$  para todo  $\xi \in \Gamma, x \in \Omega$  [32].

Notamos pela definição que a existência da Função de Green está diretamente relacionada à existência de solução da Equação de Laplace com a condição de Dirichlet dada pela função de singularidade  $s(\cdot, \cdot)$  na fronteira do domínio  $\Omega$ , como na observação abaixo: *Observação:* A função de Green existe se e somente se para todo  $x \in \Omega$  fixado o problema de valor de contorno [12].

$$\begin{cases} \nabla^2 \Phi = 0 \text{ em } \Omega \\ \Phi = -s(\cdot, x) \text{ sobre } \Gamma \end{cases} \quad (2.17)$$

tem solução  $\Phi$  em  $C^2(\overline{\Omega})$ .

Como apresentado em [12] se  $\Phi$  é solução do problema de valor de contorno acima, então a função de Green para a região  $\Omega$  é

$$G(x, y) = s(x, y) - \Phi \quad x, y \in \Omega, x \neq y \quad (2.18)$$

**Teorema 2.1.7.** Seja  $\Omega$  um domínio normal. Seja  $\nabla^2 u = f$  e a condição de Dirichlet com solução  $u \in C^2(\overline{\Omega})$ . Assumindo a existência de uma função de Green de primeiro tipo. Então pode-se expressar  $u$  explicitamente por

$$u(x) = - \int_{\Omega} G(\xi, x) f(\xi) d\xi - \int_{\Gamma} \phi(\xi) \frac{\partial G(\xi, x)}{\partial n} d\Gamma_{\xi} \quad (2.19)$$

## 2.1.5 Método de Equações Integrais

Como visto anteriormente, a função de singularidade  $s$  é uma solução fundamental da equação de Laplace. Desse modo, se consideramos tal solução  $s$  em lugar de  $\gamma$  na representação (2.16), e considerando adicionalmente que impomos a condição de Neumann  $\partial u / \partial n = \phi$  na fronteira, obtemos:

$$u(x) = - \int_{\Gamma} \frac{\partial s}{\partial n_{\xi}} u(\xi) d\Gamma_{\xi} + \int_{\Gamma} s(\xi, x) \phi(\xi) d\Gamma_{\xi} - \int_{\Omega} s(\xi, x) f(\xi) d\Gamma_{\xi} \quad (2.20)$$

Denotando a função núcleo  $k(x, \cdot) := -\partial s(\xi, x) / \partial n_{\xi}$  e  $G(x) := G_1(x) + G_2(x)$  onde

$$\begin{aligned} G_1 &:= \int_{\Gamma} s(\xi, x) \phi(\xi) d\Gamma_{\xi} \\ G_2 &:= - \int_{\Omega} s(\xi, x) f(\xi) d\Gamma_{\xi} \end{aligned} \quad (2.21)$$

podemos reescrever  $u$  como

$$u(x) = \int_{\Gamma} k(x, \xi) u(\xi) d\Gamma_{\xi} + G(x) \quad \forall x \in \Omega. \quad (2.22)$$

O primeiro termo da direita de (2.22) é chamado potencial de camada dupla (potencial de dipolo),  $G_1$  é o potencial de camada simples e  $G_2$  é o potencial de volume.

### 2.1.6 Modelos Contínuos por Formulações integrais

A partir do Teorema 2.1.4 vimos que é possível determinar uma fórmula para  $u(\mathbf{y}) \quad \forall \mathbf{y} \in \mathbb{R}^3 \setminus \{\Gamma\}$  se soubermos  $u(\mathbf{x})$  e  $\frac{\partial u(\mathbf{x})}{\partial n} \quad \forall \mathbf{x} \in \Gamma$ .

Consideremos uma função  $W : \mathbb{R}^3 \rightarrow \mathbb{R}$  tal que

$$\begin{cases} -\nabla^2 W = 0 & \text{em } \mathbb{R}^3 \setminus \overline{\Omega} \\ W \rightarrow 0 & \text{no infinito} \end{cases}$$

Seja  $\mathbf{n}(\mathbf{y})$  o vetor normal em  $\mathbf{y} \in \Gamma$  e suponhamos que os seguintes limites existam para todo  $x \in \Omega$ :

$$W_i(\mathbf{x}) = \lim_{\eta \rightarrow 0^+} W(\mathbf{x} - \eta \mathbf{n}(\mathbf{x})); \quad \left. \frac{\partial W}{\partial n} \right|_i = \lim_{\eta \rightarrow 0^+} \frac{W_i(\mathbf{x}) - W_i(\mathbf{x} - \eta \mathbf{n}(\mathbf{x}))}{\eta} \quad (2.23)$$

$$W_e(\mathbf{x}) = \lim_{\eta \rightarrow 0^+} W(\mathbf{x} + \eta \mathbf{n}(\mathbf{x})); \quad \left. \frac{\partial W}{\partial n} \right|_e = \lim_{\eta \rightarrow 0^+} \frac{W_e(\mathbf{x}) - W_e(\mathbf{x} + \eta \mathbf{n}(\mathbf{x}))}{\eta} \quad (2.24)$$

Desse modo podemos definir, para  $\mathbf{x} \in \Gamma$ , os saltos:

$$[W](\mathbf{x}) = W_i(\mathbf{x}) - W_e(\mathbf{x}) \quad (2.25)$$

$$\left[ \frac{\partial W}{\partial n} \right](\mathbf{x}) = \left. \frac{\partial W}{\partial n} \right|_i - \left. \frac{\partial W}{\partial n} \right|_e \quad (2.26)$$

e garantir que sejam bem definidos para todo  $\mathbf{x} \in \Gamma$ .

A existência desses limites permite, a partir do valor principal de Cauchy [13], que usemos o teorema 2.1.4 para estabelecer uma fórmula para determinar  $W$  para todo  $\mathbf{y} \in \Omega \setminus \{\Gamma\}$ .

$$W(\mathbf{y}) = \int_{\Gamma} \frac{1}{4\pi|\mathbf{y} - \mathbf{x}|} \left[ \frac{\partial W}{\partial n} \right](\mathbf{x}) d\mathbf{x} - \int_{\Gamma} \frac{\partial}{\partial n_{\mathbf{x}}} \left( \frac{1}{4\pi|\mathbf{y} - \mathbf{x}|} \right) [W](\mathbf{x}) d\mathbf{x} \quad (2.27)$$

A função  $W$  não tem um valor bem definido sobre  $\Gamma$ . Por outro lado, o fato de (2.27) depender da norma entre  $\mathbf{y} \in \Omega \setminus \{\Gamma\}$  e  $x \in \Gamma$  e pelos limites (2.23) e (2.24) são válidas as seguintes relações para todo  $x^* \in \Gamma$ :

$$\frac{W_i(\mathbf{x}^*) + W_e(\mathbf{x}^*)}{2} = \int_{\Gamma} \frac{1}{4\pi|\mathbf{x}^* - \mathbf{x}|} \left[ \frac{\partial W}{\partial n} \right] (\mathbf{x}) d\mathbf{x} - \int_{\Gamma} \frac{\partial}{\partial n_{\mathbf{x}}} \left( \frac{1}{4\pi|\mathbf{x}^* - \mathbf{x}|} \right) [W](\mathbf{x}) d\mathbf{x} \quad (2.28)$$

$$\frac{1}{2} \left( \frac{\partial W}{\partial n} \Big|_i + \frac{\partial W}{\partial n} \Big|_e \right) (\mathbf{x}^*) = \int_{\Gamma} \frac{\partial}{\partial n_{\mathbf{x}}} \left( \frac{1}{4\pi|\mathbf{x}^* - \mathbf{x}|} \right) \left[ \frac{\partial W}{\partial n} \right] (\mathbf{x}) d\mathbf{x} - \int_{\Gamma} \frac{\partial^2}{\partial n_{\mathbf{x}} \partial n_{\mathbf{x}^*}} \left( \frac{1}{4\pi|\mathbf{x}^* - \mathbf{x}|} \right) [W](\mathbf{x}) d\mathbf{x}. \quad (2.29)$$

As fórmulas de representação integral sugerem a introdução dos operadores  $S$ ,  $D$ ,  $D^*$  e  $N$  definidos para  $\sigma : \Gamma \rightarrow \mathbb{R}^3$  e  $\mathbf{x} \in \Gamma$  por

$$(S\sigma)(\mathbf{x}) = \int_{\Gamma} \frac{1}{|\mathbf{x} - \mathbf{x}^*|} \sigma(\mathbf{x}^*) d\mathbf{x}^* \quad (2.30)$$

$$(D\sigma)(\mathbf{x}) = \int_{\Gamma} \frac{\partial}{\partial n_{\mathbf{x}^*}} \left( \frac{1}{|\mathbf{x} - \mathbf{x}^*|} \right) \sigma(\mathbf{x}^*) d\mathbf{x}^* \quad (2.31)$$

$$(D^*\sigma)(\mathbf{x}) = \int_{\Gamma} \frac{\partial}{\partial n_{\mathbf{x}}} \left( \frac{1}{|\mathbf{x} - \mathbf{x}^*|} \right) \sigma(\mathbf{x}^*) d\mathbf{x}^* \quad (2.32)$$

$$(N\sigma)(\mathbf{x}) = \int_{\Gamma} \frac{\partial^2}{\partial n_{\mathbf{x}} \partial n_{\mathbf{x}^*}} \left( \frac{1}{|\mathbf{x} - \mathbf{x}^*|} \right) \sigma(\mathbf{x}^*) d\mathbf{x}^* \quad (2.33)$$

Quando o contorno  $\Gamma$  é regular ( $C^1$  pelo menos), os núcleos de Green dos operadores  $S$ ,  $D$ ,  $D^*$  apresentam singularidades integráveis [22]. O operador  $N$  não será utilizado nos métodos que serão tratados aqui, portanto nenhum detalhes dele será abordado.

Os operadores  $S$ ,  $D$  e  $D^*$  satisfazem as seguintes propriedades [22]:

- Propriedade 1: em  $L^2(\Gamma)$ , o operador  $S$  é auto-adjunto e  $D$  é adjunto de  $D^*$ .
- Propriedade 2:  $DS = SD^*$
- Propriedade 3: denotando por  $H^s$  o espaço de Sobolev de índice  $s \in \mathbb{R}$ , as aplicações:

$$S : H^s(\Gamma) \rightarrow H^{s+1}(\Gamma)$$

e

$$\lambda - D^* : H^s(\Gamma) \rightarrow H^s(\Gamma) \quad \text{para } -2\pi < \lambda < +\infty$$

são isomorfismos bicontínuos para todo  $s \in \mathbb{R}$ .

Para concluir essa seção matemática, destacamos que as funções potenciais de camada simples (2.21) e dupla (2.22) apresentada na seção anterior satisfazem as condições da função

$W$  assim como os limites apresentados nas equações (2.23) e (2.24) existem. Por identificação com a equação (2.27) temos

$$[W_{cs}] = 0 \quad \left[ \frac{\partial W_{cs}}{\partial n} \right] = 4\pi\sigma \quad (2.34)$$

$$[W_{cd}] = 4\pi\sigma \quad \left[ \frac{\partial W_{cd}}{\partial n} \right] = 0 \quad (2.35)$$

### 2.1.7 Campo de Energia de Reação das Cargas Interiores

O potencial de reação  $\Phi_R$  é definido como  $\Phi_R = \Phi - \Phi_M$  onde  $\Phi$  é a solução única para

$$-\nabla \cdot (\epsilon(\mathbf{x}) \nabla \Phi(x)) = 4\pi\rho_M(\mathbf{x}) \quad (2.36)$$

se anulando no infinito [8] e onde

$$\Phi_M(\mathbf{x}) = \int_{\mathbb{R}^3} \frac{\rho_M(x^*)}{|\mathbf{x} - \mathbf{x}^*|} d\mathbf{x}^* \quad (2.37)$$

denota o potencial gerado por  $\rho_M$  no vácuo. Como  $\epsilon(\mathbf{x}) = 1$  em  $C$  e  $\epsilon(\mathbf{x}) = \epsilon_0$  fora de  $C$  e como  $\rho_M$  é assumido ter suporte no interior de  $C$  temos

$$\begin{cases} -\nabla^2 \Phi = 4\pi\rho_M & \text{em } C \\ -\nabla^2 \Phi = 0 & \text{fora de } C \\ \Phi \rightarrow 0 & \text{no infinito} \end{cases}$$

Da mesma forma, o potencial  $\Phi_M$  também satisfaz

$$\begin{cases} -\nabla^2 \Phi_M = 4\pi\rho_M & \text{em } C \\ -\nabla^2 \Phi_M = 0 & \text{fora de } C \\ \Phi_M \rightarrow 0 & \text{no infinito} \end{cases}$$

Como  $\Phi_R = \Phi - \Phi_M$  temos:

$$\begin{cases} -\nabla^2 \Phi_R = 0 & \text{em } C \\ -\nabla^2 \Phi_R = 0 & \text{fora de } C \\ \Phi_R \rightarrow 0 & \text{no infinito} \end{cases}$$

Mais uma consequência do fato de  $\rho_M$  ter suporte no interior da cavidade  $C$  é que os limites dados em (2.23) e (2.24) para o potencial de reação existem. Portanto podemos usar a representação integral tipo *camada simples* (2.34) para determinar o potencial de reação para todo  $\mathbf{x} \in \mathbb{R}^3$ . Ou seja, teremos

$$\Phi_R(\mathbf{x}) = \int_{\Gamma} \frac{\sigma(\mathbf{x}^*)}{|\mathbf{x} - \mathbf{x}^*|} d\mathbf{x}^*; \quad \forall \mathbf{x} \in \mathbb{R}^3, \quad (2.38)$$

onde

$$\sigma = \frac{1}{4\pi} \left[ \frac{\partial \Phi_R}{\partial n} \right]. \quad (2.39)$$

Para se calcular a distribuição aparente de cargas  $\sigma$  sobre a superfície da cavidade  $C$  (ASC - Aparent Surface Charge) faremos uso das seguintes relações obtidas diretamente pela definição dos saltos para os potenciais *camada simples* e *camada dupla* e pelos valores de permissividade elétrica para cada meio. Estamos considerando o solvente como um dielétrico isotropico com uma permissividade  $\epsilon_x$  e a cavidade como um espaço rarefeito o suficiente para que sua permissividade seja  $\epsilon_0 = 1$ , ou seja, permissividade elétrica no vácuo.

$$\frac{\partial \Phi_R}{\partial n} \Big|_i - \frac{\partial \Phi_R}{\partial n} \Big|_e = 4\pi\sigma \quad (2.40)$$

$$\frac{1}{2} \left[ \frac{\partial \Phi_R}{\partial n} \Big|_i + \frac{\partial \Phi_R}{\partial n} \Big|_e \right] = D^* \sigma \quad (2.41)$$

$$\begin{aligned} 0 &= \frac{\partial \Phi}{\partial n} \Big|_i - \epsilon_0 \frac{\partial \Phi}{\partial n} \Big|_e \\ &= \frac{\partial \Phi_R}{\partial n} \Big|_i - \epsilon_0 \frac{\partial \Phi_R}{\partial n} \Big|_e + (1 - \epsilon_0) \frac{\partial \Phi_R}{\partial n} \end{aligned} \quad (2.42)$$

observando que em geral a distribuição de cargas  $\rho_M$  das moléculas do soluto no vácuo é conhecida e portanto é possível calcular  $\Phi_M$  e sua derivada normal partir de (2.37).

Essa considerações nos levam a equação integral

$$\left( 2\pi \frac{\epsilon + 1}{\epsilon - 1} - D^* \right) \sigma = \frac{\partial \Phi_M}{\partial n} \quad (2.43)$$

Essa é a formulação mais clássica dentre os modelos PCM e é referenciada como DPCM [17] [6]. A discussão sobre as propriedades que garantem a existência e unicidade da solução  $\sigma$  podem ser vistas em [22].

O campo de energia de reação  $E_R(\rho, \rho')$  pode ser finalmente reescrito com uma integral sobre  $\Gamma$ :

$$\begin{aligned} E_R(\rho, \rho') &= \int_{\mathbb{R}^3} \rho'(\mathbf{x}) \Phi_R(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathbb{R}^3} \rho'(\mathbf{x}) \left( \int_{\Gamma} \frac{\sigma(\mathbf{x}^*)}{|\mathbf{x} - \mathbf{x}^*|} d\mathbf{x}^* \right) d\mathbf{x} \\ &= \int_{\Gamma} \sigma(\mathbf{x}^*) \left( \int_{\mathbb{R}^3} \frac{\rho'(\mathbf{x})}{|\mathbf{x} - \mathbf{x}^*|} d\mathbf{x} \right) d\mathbf{x}^* \\ &= \int_{\Gamma} \sigma(\mathbf{x}^*) \Phi'_M(\mathbf{x}^*) d\mathbf{x}^* \end{aligned} \quad (2.44)$$

Explorando as propriedades dos operadores sobre Equação (2.43) vários modelos com formulação usando equações integral (IEF) foram construídos. Aplicando-se o operador  $S$  sobre ambos os lados da equação (2.43), usando a propriedade comutativa  $SD^* = DS$  e a fórmula de representação (2.28) é possível obter uma família de formulações integrais para se determinar a distribuição  $\sigma$  [17]. Em particular, nos interessa, para o desenvolvido desse trabalho, as formulações IEF-PCM e o COSMO.

$$\text{IEF-PCM} \quad S \left[ 2\pi \frac{\epsilon + 1}{\epsilon - 1} - D^* \right] \sigma = -\frac{\epsilon - 1}{\epsilon} (2\pi - D) \Phi_M \quad (2.45)$$

COSMO

$$\begin{cases} S\tilde{\sigma} = -\Phi_M \\ (2\pi \frac{\epsilon+1}{\epsilon-1} - D^*) \sigma = \frac{\epsilon-1}{\epsilon} (2\pi - D^*) \tilde{\sigma} \end{cases} \quad (2.46)$$

A vantagem da formulação IEF-PCM sobre a DPCM está na simetria apresentada pelo operador integral da primeira. Já o modelo COSMO (COductor-like Screen MOdel) é uma aproximação do modelo IEF-PCM assintoticamente exata no limite das constantes dielétricas  $\epsilon$  de valor alto, donde seu nome. Usando o operador  $S$  adquadamente sobre a segunda equação da formulação COSMO demonstra-se sua equivalencia com a formulação IEF-PCM.

$$\begin{aligned} S \left( 2\pi \frac{\epsilon+1}{\epsilon-1} - D^* \right) \sigma &= S \left[ \frac{\epsilon-1}{\epsilon} (2\pi - D^*) \right] \tilde{\sigma} \\ &= \left( \frac{\epsilon-1}{\epsilon} \right) S [(2\pi - D^*)] \tilde{\sigma} \\ &= \left( \frac{\epsilon-1}{\epsilon} \right) [(2\pi - D)] S \tilde{\sigma} \\ &= -\left( \frac{\epsilon-1}{\epsilon} \right) [(2\pi - D)] \Phi_M \end{aligned} \quad (2.47)$$

No COSMO efetivamente usa-se o operador  $S$  e o potencial  $\Phi_M$  para se obter de forma exata de uma distribuição superficial  $\tilde{\sigma}$ . Uma aproximação é usada na segunda parte da (2.46) que é substituída por

$$\sigma \simeq \frac{\epsilon - 1}{\epsilon + k} \tilde{\sigma}$$

onde  $k$  é um parâmetro empírico, cujos valores ótimos são relatados entre  $k = 0$  e  $k = 1/2$  [22].

## 2.1.8 Métodos de Discretização

Dada a complexidade ou impossibilidade de se resolver analiticamente equações integrais como (2.45) no caso de cavidades/distribuição de carga mais gerais faz-se natural o recurso das soluções numéricas, onde, em geral, os métodos da colocação e Galerkin são propostos [4].

Para demonstrar o funcionamento desses métodos, vamos definir uma equação cuja solução exata  $u(x) \in V$  seja conhecida e aplicar ambos os métodos para resolvê-la numeri-

camente. Essa escolha particular não tira a generalidade dos métodos, sendo justificada pela simplificação na abordagem.

Seja

$$\frac{d^2 u(x)}{dx^2} + x = 0, \text{ onde } x \in \Omega = [0, 1] \quad (2.48)$$

sujeito a condição de Dirichlet que  $u(0) = u(1) = 0$

A solução analítica da equação (2.48) é  $u_{exato}(x) = \frac{x}{6} - \frac{x^3}{6}$ .

Vamos agora definir uma aproximação  $u^*(x) \in V_a \subset V$  para  $u(x)$  tal que

$$u^*(x) = \sum_{i=1}^N \alpha_i \phi_i(x) \quad (2.49)$$

onde as funções  $\phi_i(x)$ , referenciadas como funções de interpolação, são linearmente independentes e  $\{\phi_i\}$  forma uma base para  $V_a$ . Os coeficientes  $\alpha_i$  são constantes desconhecidas, determinadas ao se forçar que  $u^*(x)$  satisfaça a equação (2.48) para uma série de valores  $x_i$  do domínio  $\Omega$ .

Não se pode garantir que  $u^*(x)$  satisfaça (2.48)  $\forall x \in [0, 1]$  ou mesmo que  $u^*(x)$  satisfaça as condições de contorno. Entretanto, podemos definir uma função de erro ou resíduo  $R(x) \forall x \in ]0, 1[$  substituindo a  $u(x)$  por  $u^*(x)$  em (2.48) e para  $x = 0$  ou  $x = 1$  temos  $R(x)$  dado pela diferença  $u^*(x) - u(x)$ . Ou seja,

$$R(x) = \begin{cases} \frac{d^2 u^*(x)}{dx^2} + x & \text{se } x \in ]0, 1[ \\ u^*(x) - u(x) & \text{se } x \in \{0, 1\} \end{cases} \quad (2.50)$$

Um método de aproximação acurado é tal que se consiga minimizar esse resíduo  $R(x)$  aplicando-se diferentes técnicas que forcem a anulação desse erro em alguns pontos, regiões ou distribuições e suavizar o erro onde  $R(x) \neq 0$ . Isto é, efetivamente faz-se uma distribuição do erro que, em geral, é feito ao se forçar que o produto interno  $\langle R, \varphi_i \rangle$  de  $R(x)$  por funções de ponderação linearmente independentes  $\varphi_i$  se anule, resultando no sistema:

$$\langle R, \varphi_i \rangle = \int_{\Omega} R \varphi_i d\Omega = 0, \quad \text{para todo } i = 1, \dots, N. \quad (2.51)$$

Em virtude disso são chamadas de técnicas de *resíduos ponderados*.

Diferentes formas de  $\varphi_i$  definem os diferentes métodos de aproximação:

1. No método da colocação, escolhemos inicialmente um conjunto de  $N$  pontos  $x_1, x_2, \dots, x_n$  no domínio  $\Omega$ . Tais pontos definem as seguintes funções de ponderação:

$$\varphi_i = \delta(x - x_i) \quad \text{para todo } i = 1, \dots, N, \quad (2.52)$$

onde  $\delta(x - x_i)$  é a função Delta de Dirac centrada em  $x_i$ , a qual tem a propriedade de

$$\int_{\Omega} f(x) \delta(x - x_i) dx = f(x_i). \quad (2.53)$$

Considerando tais funções de ponderação em (2.51) obtemos, com  $f = R$  na propriedade acima, o sistema de equações:

$$R(x_i) = 0; \quad \text{para todo } i = 1, \dots, N. \quad (2.54)$$

Esse é o método mais simples, contudo não nos permite uma boa análise do erro, uma vez que é uma equação obtida pontualmente.

2. O método de Galerkin é amplamente utilizado e é obtido considerando-se como funções de ponderação as mesmas funções utilizadas para a aproximação de  $u^*(x)$ . Desse modo, como  $\varphi_i = \phi_i$ , obtemos que o sistema de equações é:

$$\int_{\Omega} R\phi_i dx = 0; \quad \text{para todo } i = 1, \dots, N. \quad (2.55)$$

Uma questão natural é que funções podemos considerar para  $\phi_i$ . Uma possibilidade é tomar funções polinomiais por partes, ou seja, dados  $k$  intervalos disjuntos do domínio  $\Omega$ , assumir que  $\phi_i$  tenha suporte compacto em  $[x_i, x_{i+1}]$ . O caso mais simples se dá quando assumimos que as funções de interpolação são constantes, sendo necessário apenas um ponto de amostragem por intervalo para se calcular os coeficientes  $\alpha$  e assim determinar a aproximação  $u^*(x)$ . Cada intervalo  $[x_i, x_{i+1}]$  do domínio em conjunto com sua  $\phi_i$  e seus pontos de amostragem (nós)  $x_j \in [x_i, x_{i+1}]$  formam um *elemento*, sendo um componente básico dos métodos conhecidos como Elementos Finitos [15] ou Elementos de Contorno [4].

Como ilustração, vamos obter uma aproximação para a equação (2.48) fazendo uso dos métodos da colocação e Galerkin. Vamos assumir que:

$$u^*(x) = \alpha_1 \phi_1(x) + \alpha_2 \phi_2(x) \quad (2.56)$$

onde

$$\phi_1 = x - 2x^2 + x^3 \quad \text{e} \quad \phi_2 = x^3 - x^2 \quad (2.57)$$

e os pontos  $x_1 = 0.25$  e  $x_2 = 0.75$  como nós para o método da colocação.

Resolvendo os sistemas (2.54) ou (2.55) encontramos  $\alpha_1 = \frac{1}{6}$  e  $\alpha_2 = -\frac{1}{3}$ . Logo, para esse caso em particular, a função de aproximação  $u^*(x) = u_{exata}(x)$ .

Em geral, metodologias por elementos são caracterizadas por dois processos:

- A definição de uma malha ou tesselação para  $\Omega$ . Cada elemento ou tessera da malha pode ter a forma de triângulos planos ou curvos, quadrados planos ou curvos ou mesmo outros

polígonos curvos.

- A escolha das funções  $\phi_i$  que formam uma base para  $V_a$ , que podem ser definidas baseadas em diversos critérios [15].

Consideremos agora uma equação integral  $A\sigma = g$ , onde  $\sigma \in H^s(\Gamma)$ ,  $g \in H^{s*}(\Gamma)$   $A \in \mathcal{L}(H^s(\Gamma), H^{s*}(\Gamma))$  é um operador integral caracterizado por um núcleo de Green  $k_A$ :

$$(A\sigma)(x^*) = \int_{\Gamma} k_A(x, x^*)\sigma(x^*)dx^*, \quad \forall x^* \in \Gamma. \quad (2.58)$$

Baseado nos processos acima, para se obter uma aproximação para  $\sigma$  é necessário determinar uma tesselação sobre  $\Gamma$  e escolher as funções  $\phi_i$  apropriadas.

No procedimento padrão descrito nos modelos PCM,  $\Gamma$  representa a superfície da cavidade molecular e é tesselado por  $K$  tesseras triangulares ( $T_k$ ) com  $1 \leq k \leq K$ , onde para o cálculo das quadraturas, é necessário, para cada elemento triangular da superfície, sua área, o seu 'centro' e o vetor normal neste centro. A distribuição de cargas  $\sigma$  é aproximada em  $V_a$  por funções constantes por partes (por tessera), onde pelo método de Galerkin  $\sigma$  é o elemento de  $V_a$  que satisfaz

$$\int_{\Gamma} \left( \int_{\Gamma} k_A(x, x^*)\sigma(x^*)dx^* \right) \varphi_i(x)dx = \int_{\Gamma} g(x)\varphi_i(x)dx, \quad \forall \varphi_i \in V_a \quad (2.59)$$

Obtem-se assim um sistema de equações de ordem  $K$

$$\mathbf{A} \cdot \sigma = \mathbf{g} \quad (2.60)$$

onde  $\sigma$  é um vetor coluna contendo o valor aproximado da densidade sobre uma única tessera,  $\mathbf{A}$  é uma matriz  $K \times K$  e  $\mathbf{g}$  é um vetor coluna que depende da distribuição de cargas  $\rho$

$$[A]_{ij} = \int_{T_i} \left( \int_{T_j} k_A(x, x^*)\sigma(x^*)dx^* \right) dx, \quad [g]_i = \int_{T_i} g. \quad (2.61)$$

O uso dessa aproximação torna simples o cálculo dos coeficientes da matriz formada quando tomamos  $A$  o operador em (2.43).

$$\begin{aligned} [A]_{ij} &= \left[ 2\pi \frac{\epsilon + 1}{\epsilon - 1} - D^* \right]_{ij} \\ &= 2\pi \frac{\epsilon + 1}{\epsilon - 1} A_i A_j - [D^*]_{ij} \\ &= 2\pi \frac{\epsilon + 1}{\epsilon - 1} A_i A_j - \int_{T_j} \left( \int_{T_i} \frac{\partial}{\partial n} \left( \frac{1}{|x - x^*|} \right) dx^* \right) dx \end{aligned}$$

Podemos agora efetivamente elaborar um algoritmo para resolver, por exemplo, a (2.43). Esse mesmo algoritmo é válido para as outras equações da família PCM.

1. Gerar uma tesselação para a superfície  $\Gamma$  da cavidade.
2. Calcular os elementos da matriz

$$\left[ 2\pi \frac{\epsilon + 1}{\epsilon - 1} - D^* \right]_{ij}$$

3. Determinar os elementos da matriz

$$[g]_i = \int_{T_i} \frac{\partial \Phi_M}{\partial n}$$

4. Resolver o sistema linear

$$[A][\sigma] = [g]$$

5. Calcular  $E_R(\rho, \rho^*)$  pela aproximação

$$E_R(\rho, \rho^*) \simeq E_R^{approx}(\rho, \rho^*) = \sum_{i=1}^N [\sigma]_i \int_{T_i} \Phi_M$$

onde a integral é calculada numericamente.

Como comentário adicional, temos que quando as densidades de carga  $\rho$  e  $\rho^*$  são compostas por cargas pontuais, dipolos ou funções gaussianas/polinomiais, o potencial  $V_M$  e sua derivada direcional  $\frac{\partial V_M}{\partial n}$  podem ser dados analiticamente.

### 2.1.9 Aspectos Numéricos

Há alguns comentários importantes sobre as integrais numéricas realizadas no algoritmo descrito acima.

Em primeiro lugar, vamos assumir que todas as integrais numéricas sejam feitas com quadraturas de ponto médio.

$$\int_{\Gamma} \lambda(x, y) dx dy = f(\alpha, \beta) A_{\Gamma}, \quad (2.62)$$

e que

$$\frac{\partial}{\partial n} \left( \frac{1}{|x - x^*|} \right) = \nabla_x \left( \frac{1}{|x - x^*|} \right) \cdot n(x) = \frac{(x - x^*) \cdot n(x)}{|x - x^*|^3} \quad (2.63)$$

No próximo capítulo serão apresentadas as formas efetivas da cavidade molecular  $C$  e consequentemente seu contorno  $\Gamma$ . É importante adiantar que, nos geradores de malha clássicos,  $\Gamma$  é obtido pela superposição de esferas [25, 7, 30]. Portanto, todos os cálculos apresentados

a seguir são feitos para triângulos esféricos, ou seja, cada tessera tem a forma de um triângulo sobre uma esfera. Na abordagem MultiPatch, admitem-se também tesseras triangulares não planas sobre toros.

A partir do exposto, obtemos explicitamente

$$\int_{T_i} \int_{T_j} dx^* dx = A_i A_j \quad (2.64)$$

onde  $A_i$  e  $A_j$  são as áreas das tesseras  $T_i$  e  $T_j$  respectivamente.

De igual forma, o operador *camada simples*  $S$  se torna

$$S_{ij} = \int_{T_i} \int_{T_j} \frac{1}{|x - x^*|} dx^* dx = A_i A_j \frac{1}{|C_i - C_j|} \quad (2.65)$$

onde  $C_i$  e  $C_j$  são os 'centros' ou pontos de amostragem sobre as tesseras  $T_i$  e  $T_j$ . Essa expressão apresenta uma singularidade quando  $i = j$  que pode ser calculada no sentido dado pelo valor principal de Cauchy. Por [19] temos

$$S_{ii} = \int_{T_i} \int_{T_i} \frac{1}{|x - x^*|} dx^* dx = 1.0694 \sqrt{\frac{4\pi}{A_i}} \quad (2.66)$$

Para o operador *camada dupla*  $D$  temos

$$D_{ij} = \int_{T_j} \left( \int_{T_i} \frac{\partial}{\partial n} \left( \frac{1}{|x - x^*|} \right) dx^* \right) dx = \frac{(C_i - C_j) \cdot n(j)}{|C_i - C_j|^3} \quad (2.67)$$

Outra singularidade aparece para  $i = j$  no operador  $D$ , que passa a ser calculada por [17]

$$D_{ii} = 1.0694 \frac{\sqrt{4\pi A_i}}{2R_i} \quad (2.68)$$

onde  $R_i$  é o raio da esfera que contem a tessera  $T_i$ .

## CAPÍTULO 3

---

### DISCRETIZAÇÃO DA SES

---

Para realização dos cálculos propostos nos modelos da família PCM faz-se necessário definir os limites entre o solvente e o soluto definindo-se uma forma geométrica para a molécula do soluto, onde, em geral, se assume como referência que cada átomo ou grupo de átomos que compoem a molécula do soluto esteja associado a uma esfera de centro  $C_{A_i} = (x_{A_i}, y_{A_i}, z_{A_i})$  e raio  $R_{A_i}$  e algumas grandezas relativas ao solvente.

Chama-se cavidade molecular  $C$  a região do espaço ocupado pela molécula de soluto e de superfície molecular ao bordo  $\Gamma = \partial C$  dessa cavidade. A forma específica da cavidade/superfície está sujeita ao modelo e a componente energética, apresentando em alguns casos maior simetria e em outros mais “realismo”. Vejamos três definições mais frequentes [22]:

1. A superfície de van der Waals (VWS) é definida como a superfície obtida por um conjunto de esferas sobrepostas, cada qual centrada em um átomo ou grupo de átomos e tendo como raio o correspondente raio de van der Waals[7]. A VWS é comumente usada para calcular a contribuição da cavitação para a energia livre de solvatação  $\Delta G_s^*$ .
2. A superfície acessível do solvente (SAS) é definida como a superfície determinada pelo conjunto de pontos descritos pelo centro de uma prova esférica do solvente rolando sobre a VWS: o raio da prova do solvente está relacionado à dimensão e à natureza do solvente. A SAS é comumente usada para calcular as contribuições de curto alcance (dispersiva e repulsiva) para a energia livre de solvatação.
3. A superfície excludente do solvente (SES) é definida como a superfície formada pelo conjunto de pontos tangentes a uma prova esférica do solvente rolando sobre a VWS. Essa superfície delimita a região do espaço onde a prova esférica do solvente não pode entrar sem que intercepte a VWS. A SES se aparenta como a VWS na qual as fendas da

interseção entre esferas é suavizada; a parte convexa da SES é compartilhada com a VWS e é chamada de superfície de contato, enquanto que a parte que não é compartilhada com a VWS é chamada de superfície reentrante. A SES é comumente usada para calcular a contribuição eletrostática para a energia livre de solvatação. A região do espaço que é incluída na SES e não incluída na VWS é chamada de volume excludente do solvente.

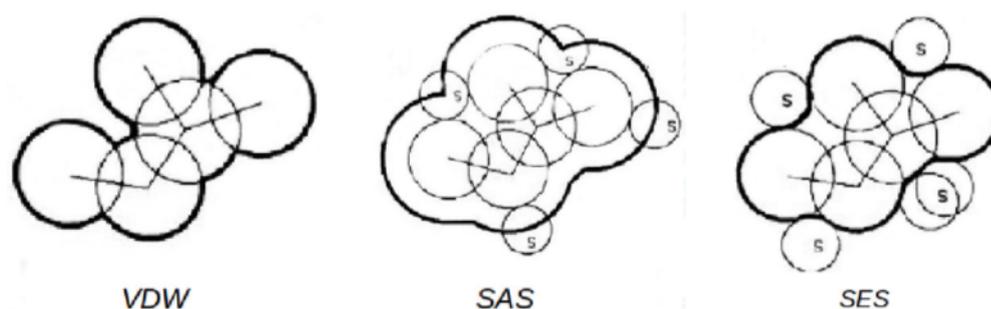


Figura 3.1: Nessa figura as bordas mais espessas representam a superfície da cavidade quando descrita por esferas de van der Walls (VDW), superfície acessível ao solvente (SAS), superfície excludente do solvente (SES).

Para realização dos cálculos numéricos a cavidade deve ser tessalada. Uma tesselação é uma partição de uma superfície em subregiões chamadas tesseras, sendo tradicionalmente elementos de área triangulares sobre esferas. Como conhecido em outras áreas, uma tesselação da superfície é o mesmo que uma malha sobre a superfície.

Uma tesselação para as duas primeiras superfícies pode ser facilmente obtida a partir de um conjunto de esferas sobrepostas. O mesmo não vale para a SES. Existem vários algoritmos [25, 7, 30] que buscam traduzir a definição abstrata da SES como a superfície de um sólido complexo composto por objetos geométricos simples a partir dos quais essa superfície possa ser tessalada. Uma vez que um dos objetivos é realizar uma otimização na geometria molecular, é de grande interesse que a tesselação seja diferenciável.

Uma tesselação é dita diferenciável quando é possível calcular analiticamente as derivadas com respeito a geometria molecular e é bem definida quando as grandezas relacionadas a tesselação e suas derivadas são estáveis do ponto de vista numérico [22].

### 3.1 GEPOL

O método conhecido por GEPOL [25], sigla para GEometria POLiedro, o mais amplamente utilizado, foi inicialmente elaborado por Tomasi e Pascual-Ahuir e outras melhorias foram introduzidas ao longo dos anos com o propósito de obter melhores resultados na discretização [30, 25, 20].

Nesse método, o volume excludente é aproximado por um conjunto de esferas. Mais precisamente, além das esferas referentes aos átomos da molécula do soluto, o método adiciona

outras esferas através de um algoritmo recursivo, com o objetivo de assim preencher o volume dessa cavidade. Desse modo, as esferas centradas no átomos constituem uma primeira geração de esferas. Quando  $r_{ab} < R_A + R_B + 2R_S$  onde  $r_{ab}$  é a distância entre os centros das esferas e  $R_A, R_B$  e  $R_S$  são os raios das esferas relativas ao átomo  $A, B$  e ao solvente  $S$  respectivamente, uma ou mais esferas são adicionadas. O centro das novas esferas está localizado sobre o segmento de reta que liga os centros de duas esferas geradas e a posição e o raio das esferas são escolhidos de maneira a maximizar o preenchimento do volume excluído do solvente. Esse processo é repetido recursivamente utilizando as esferas recém geradas. Alguns limites são determinados para que o processo não produza um loop infinito.

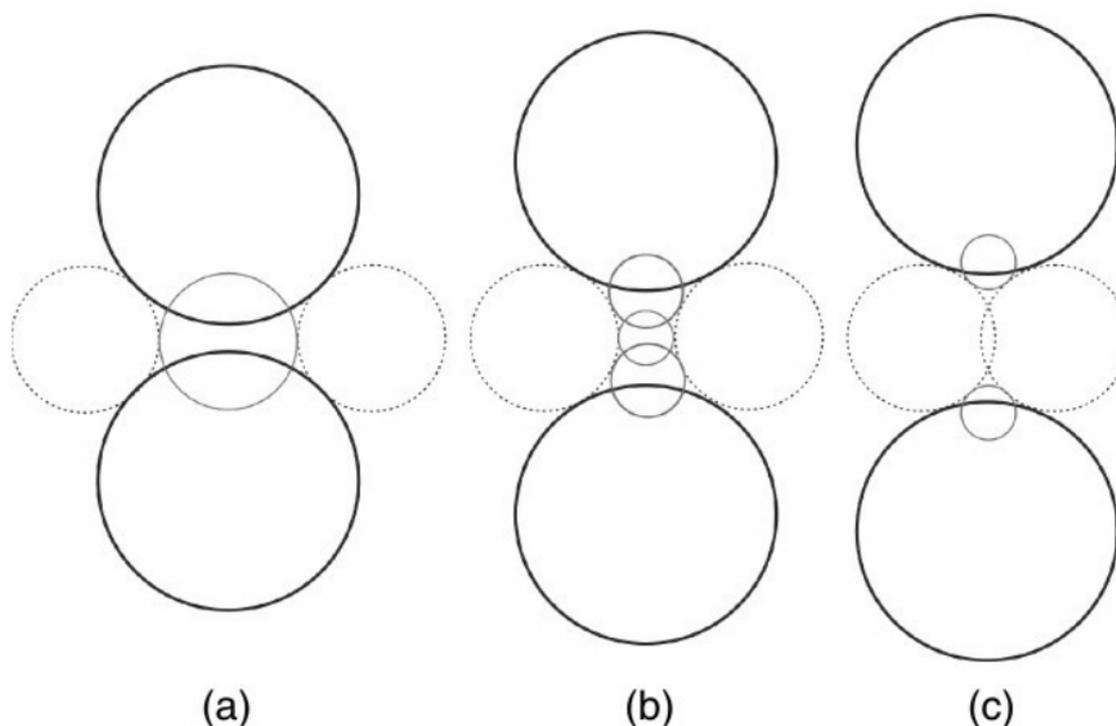


Figura 3.2: Processo de adição de esferas do GEPOL. (a) Para duas esferas próximas uma única esfera interceptando as esferas pais é gerada. (b) Para esferas mais distantes, primeiramente uma esfera que não intercepta as esferas 'pais' é gerada, então duas esferas de terceira geração são adicionadas. (c) Para qualquer par de esferas com afastamento maior, esferas menores bastante sobrepostas com as primitivas são geradas. As esferas pontilhadas representam as esferas de prova do solvente.

Pacotes quanto-mecânicos usam a simetria molecular com o objetivo de reduzir o esforço computacional. Esse comportamento pode ser utilizado se o ponto de amostragem da superfície da cavidade respeitar a simetria da molécula. Uma maneira de se satisfazer essa exigência é particionar cada superfície esférica pela respectiva simetria molecular: isso pode ser obtido pelo uso de poliedros básicos que subtenda o mesmo grupo de pontos da molécula, de maneira que a cavidade resultante seja invariante sobre qualquer transformação geométrica que faça parte do grupo de simetria molecular.

Portanto, nesse método as esferas são tesselladas inicialmente por um pentakisdo-

cahedro [9], definindo uma tesselação com 60 tesseras com mesma área. Um esquema de refinamento flexível foi introduzido usando alguns poliedros básicos, no qual as faces triangulares originais desse poliedro são particionadas através de um procedimento de divisão. O procedimento da divisão equilátera substitui cada triângulo original do poliedro por  $N^2$  triângulos, sendo  $N$  a ordem da divisão equilátera. Sendo  $M$  o número de faces do poliedro inicial, o final terá  $MN^2$  faces.

Vale notar que para a discretização da equação (2.43) por BEM [4], uma tesselação da SES é definida como exposto. Considerando que o espaço de funções da discretização é constante por elemento (tessera), podemos considerar para o cálculo das integrais, quadratura de um ponto. Desse modo, basta ao algoritmo de tesselação definir, para cada tessera, seu centro, área e vetor normal.

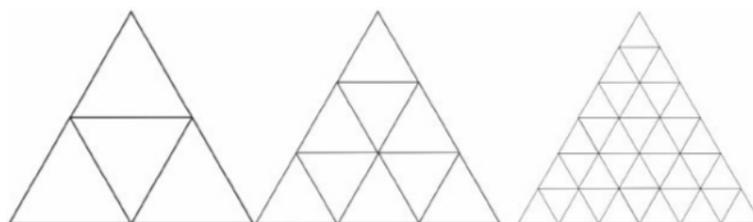


Figura 3.3: Processo de divisão equilátera

### 3.1.1 Partição das interseções entre esferas

Quando duas ou mais esferas se interceptam, algumas de suas tesseras são cortadas para excluir a porção das superfícies que estão dentro de outra esfera. No GEPOL, esse procedimento de corte testa se uma tessera intercepta uma superfície esférica e corta a parte da tessera que está dentro dela, assim, para cada tessera-interseção esférica, uma parte da tessera é cortada. Se a tessera inteira está situada dentro da esfera, será completamente removida da lista de tesseras. Tal procedimento é repetido para cada par esfera-tessera. O custo computacional desse passo pode ser reduzido, assim como feito para a generalização da adição de esferas, se uma lista de esferas próximas for previamente gerada. A primeira versão do esquema de cortes no GEPOL foi baseado em uma simples partição em sub-tesseras. A tesselação resultante não era diferenciável. Como uma tesselação diferenciável é essencial para o processo de otimização baseado no cálculo de gradientes, um cálculo analítico dos cortes das tesseras foi introduzido.

### 3.1.2 Dificuldades

Os vários passos da implementação do GEPOL que foram descritos não são completamente confiáveis do ponto de vista numérico, especialmente quando usado para a otimização de geometria baseada no cálculo de gradientes [22, 7]. A contribuição dos elementos de superfície sobre os gradientes é calculada considerando a variação da forma e área de cada tessera com

respeito ao deslocamento das esferas que se inteceptam. As esferas primitivas são centradas nos átomos e então se inicia a evolução da geometria molecular. O deslocamento das esferas adicionais está relacionado aos átomos que aparecem nas suas árvores genealógicas, assim como os raios das esferas adicionais são variáveis por definição. A evolução das esferas adicionais durante a otimização da geometria pode levar a sua eliminação quando se sobrepõem e/ou seu raio é menor que um raio de tolerância. Como resultado, essa variação do conjunto de esferas adicionais leva a uma descontinuidade na descrição do campo de reação do solvente. Casos típicos em que essa descontinuidade pode ocorrer são aqueles em que existe uma variação da distância entre dois átomos. Esse tipo de instabilidade numérica não altera o ponto estacionário final alcançado pelo procedimento de otimização, mas isso pode aumentar o número de passos necessários para alcançar esse ponto. Em casos extremos, o algoritmo de busca pode entrar em um loop infinito, na qual a geometria molecular circula sobre o ponto estacionário mas nunca o atinge.

## 3.2 DefPol

O DefPol é uma nova estratégia baseada no GEPOL, proposta por Pomelli e Tomasi [7], cujo objetivos iniciais eram:

1) Evitar o número de excessivo de parâmetros gerados pelo GEPOL, principalmente no caso de moléculas grandes com geometria irregular: muitas esferas adicionais para formação da SES e, conseqüentemente, muitas tesseras.

2) Obter sempre tesseras triangulares. Isso facilita o cálculo de derivadas analíticas necessárias para otimização de geometria.

3) Monitorar o número de tesseras geradas.

4) Evitar a ocorrência de 'bolsas de solvatação' falsas: moléculas grandes podem apresentar espaços vazios grandes o suficiente para acomodar uma molécula do solvente. O GEPOL não tem a capacidade de identificar esses casos e acaba por preencher esse espaço vazio, introduzindo uma falsa contribuição dessa região à energia livre de solvatação.

No algoritmo do DefPol, as cavidades são obtidas inicialmente a partir de uma função de forma determinada pela união de funções de forma de esferas sólidas centradas nos átomos ou grupo de átomos da molécula de soluto  $M$ :

$$f_{hs}(M) = \bigvee_G f_{hs}(G) \quad G \in M \quad (3.1)$$

onde  $\bigvee$  é o conector lógico *ou* e  $G$  é cada átomo ou grupo de átomos de  $M$ .

Para a superfície de van der Waals, temos:

$$f_w(\mathbf{p}) = \begin{cases} 1 & \text{se } |\mathbf{p} - \mathbf{p}_i| \leq R_i \\ 0 & \text{caso contrário} \end{cases} \quad (3.2)$$

Para a SAS, temos:

$$f_w(\mathbf{p}) = \begin{cases} 1 & \text{se } |\mathbf{p} - \mathbf{p}_i| \leq R_i + R_S \\ 0 & \text{caso contrário} \end{cases} \quad (3.3)$$

Para a região monoatômica da SAS, temos:

$$f_w(\mathbf{p}) = \begin{cases} 1 & \text{se } |\mathbf{p} - \mathbf{p}_i| \leq \lambda_i R_i \\ 0 & \text{caso contrário} \end{cases} \quad (3.4)$$

onde  $\lambda_i$  é um fator de escala geralmente dentro do intervalo 0.90-1.25. A região multiatômica leva em consideração a interação entre o solvente e dois ou mais átomos de soluto.

O procedimento para construção da superfície molecular com o DefPol consiste em aplicar sucessivas deformações em um poliedro inscrito em uma esfera. O algoritmo para esse procedimento é:

1. construção do poliedro;
2. projeção sobre um elipsoide de inércia;
3. deslocamento dos vértices;
4. deslocamento dos centros;
5. refinamento; e
6. tesselação

Em [7] é feita a comparação desse método com o GEPOL para 80 tipos diferentes de moléculas revela uma correlação  $r_1 = 0.999912$  para a SES e  $r_1 = 0.999990$  para a cálculo da energia livre de solvatação [7]. Além disso o método é rápido do ponto de vista numérico, mas ainda é afetado por sérios problemas numéricos ao se calcular os termos das derivadas e ao ser aplicado em moléculas com forma oblonga ou não convexa.

### 3.3 Conclusões

O estudo dos métodos GEPOL, DefPol e outras não comentadas [10, 30] sinalizou diversas questões para a elaboração de uma nova estratégia para a tesselação da SES:

- tesselação diferenciável;
- criação de tesseras sempre triangulares;
- eficiência quanto a otimização da geometria molecular.

Uma das propostas para essa nova estratégia é não utilizar apenas esferas sobrepostas para preencher o volume excludente. Levando-se em consideração que dois átomos vizinhos, cuja distância entre os centros é menor que a soma dos seus raios com o raio do solvente de prova podem estar conectados por fragmentos de toro, é possível evitar o preenchimento dessa região por esferas usando o próprio toro como referência e determinando analiticamente essa superfície, que pode ser tessellada de maneira independente da tesselação das esferas. Fragmentos esféricos concavos são utilizados para se obter a tesselação das regiões que fazem parte da SES mas não podem ser obtidas a partir das esferas 'dos átomos' ou dos fragmentos de toro.

Outra questão é a melhoria na descrição da área da superfície excludente e consequentemente do volume. Uma melhor descrição da área pode possibilitar uma menor quantidade de tesselas, o que implica na redução no custo computacional.

O principal tópico, porém, é a busca por uma melhor descrição das derivadas. Espera-se que, uma vez tendo o vetor normal correto na superfície excludente, possam ser resolvidas as questões de instabilidade numérica quanto ao cálculo das derivadas, importante para o processo de otimização da geometria molecular, uma vez que é baseado no cálculo de gradientes.

Pautado no exposto tornou-se possível a criação do método proposto por esse trabalho, cuja descrição detalhada e os resultados obtidos encontram-se nos próximos capítulos.

# CAPÍTULO 4

## MultiPatch

### 4.1 Tesselação dos objetos básicos

#### 4.1.1 Esfera

Uma parametrização  $\varphi(\phi, \theta)$  da superfície  $\Gamma$  de uma esfera de raio  $r$  pode ser dada por:

$$\varphi = (x_c + r \sin(\theta) \cos(\phi), y_c + r \sin(\theta) \sin(\phi), z_c + r \cos(\theta)) \quad (4.1)$$

onde  $0 \leq \phi \leq 2\pi$  e  $0 \leq \theta \leq \pi$ .

Para tesselação de  $\Gamma$ , considere as seguintes definições:

**Definição 4.1.1.** *Os vértices de um dodecaedro regular particular centrado na origem podem ser dado em coordenadas cartesianas por:*

$$(\pm 1, \pm 1, \pm 1), \quad (0, \pm \frac{1}{\rho}, \pm \rho), (\pm \frac{1}{\rho}, \pm \rho, 0), \quad (\pm \rho, 0, \pm \frac{1}{\rho})$$

onde  $\rho = \frac{1+\sqrt{5}}{2}$ . O comprimento da aresta é  $a = \frac{2}{\rho}$  e está inscrito numa esfera  $\Omega_0$  com centro na origem e raio  $\sqrt{3}$ .

Medindo as distâncias provamos que todos os vértices estão sobre  $\Omega_0$ . Esse vínculo permite que esse dodecaedro particular possa ser utilizado para representar qualquer dodecaedro.

**Definição 4.1.2.** *Um pentakisdodecaedro é um sólido obtido a partir de um dodecaedro. Sobre cada face do dodecaedro se constroi uma pirâmide cujo vértice é a projeção do centro da circunferência inscrita nessa face sobre  $\Omega_0$ . Esse sólido apresenta 60 faces formadas por triângulos isósceles, 32 vértices e 90 arestas.*

Um método simples para o refinamento dessa tesselação consiste em dividir cada face do poliedro em  $n^2$  triângulos equiláteros, projetando os novos vértices também sobre  $\Gamma_0$ . Esse mecanismo é chamado de divisão equilátera e  $n$  é a ordem dessa divisão.

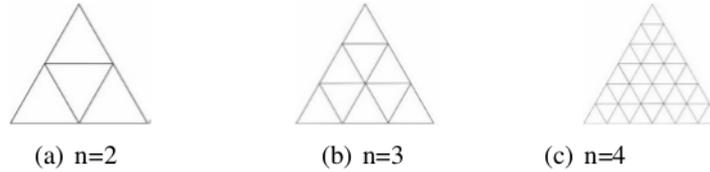


Figura 4.1: Divisão equilátera

Como as tesseras tem a mesma área, basta saber o número total de triângulos por face,  $n_{trian}$ , multiplicar pelo número de faces e dividir pela área total da esfera de raio  $r$  para determinar  $A_{TS}$ . De fato, temos:

$$\begin{aligned} n_{trian} &= 4^{n-1} \\ fndiv &= 60n_{trian} \\ A_{TS} &= \frac{4\pi r^2}{fndiv} \end{aligned} \quad (4.2)$$

O centro de cada tessera será o ponto médio entre os três pontos que definem cada triângulo sobre o pentakisdodecahedro ou sobre o triângulos definidos pela divisão equilátera. A conexão entre os vértices definida é chamada de conectividade. A conectividade é usada também para representar graficamente as tesseras.

Esse ponto médio é então projetado sobre a esfera de raio 1 centrada na origem. Assim, dados os pontos  $P_1 = (X_{V1}, Y_{V1}, Z_{V1})$ ,  $P_2 = (X_{V2}, Y_{V2}, Z_{V2})$  e  $P_3 = (X_{V3}, Y_{V3}, Z_{V3})$ , temos:

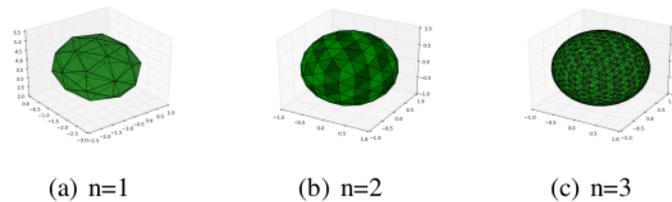


Figura 4.2: Refinamento das esferas

$$X_C = \frac{X_{V1} + X_{V2} + X_{V3}}{3} \quad (4.3)$$

$$Y_C = \frac{Y_{V1} + Y_{V2} + Y_{V3}}{3} \quad (4.4)$$

$$Z_C = \frac{Z_{V1} + Z_{V2} + Z_{V3}}{3} \quad (4.5)$$

$$R_C = \sqrt{(X_C^2 + Y_C^2 + Z_C^2)} \quad (4.6)$$

$$C = \left( \frac{X_C}{R_C}, \frac{Y_C}{R_C}, \frac{Z_C}{R_C} \right) \quad (4.7)$$

$$(4.8)$$

Se  $\varphi(\theta, \phi)$  é uma parametrização de uma superfície  $S$ , o vetor

$$\vec{N} = \frac{\partial \varphi}{\partial \theta}(\theta_0, \phi_0) \times \frac{\partial \varphi}{\partial \phi}(\theta_0, \phi_0) \neq \vec{0}$$

é um vetor normal a  $S$  em  $\varphi(\theta_0, \phi_0)$ .

Em particular para a superfície  $\Gamma$  de uma esfera de raio  $r$  e centro na origem, temos que  $\hat{n} = \frac{(x, y, z)}{r}$  é um vetor normal unitário à  $\Gamma$  no ponto  $(x, y, z)$ .

Para o processo de tesselação nos interessam os vetores normais a  $\Gamma$  no centro de cada região triangular. Assim, para cada tessera  $T_i$  de centro  $C_i$  sobre  $\Gamma$  cujo centro é  $C_\Gamma$ , temos:

$$\hat{n} = \frac{C_i - C_\Gamma}{\|C_i - C_\Gamma\|} \quad (4.9)$$

Como parâmetro adicional, podemos atribuir um identificador numérico  $i$  para cada esfera gerada por esse processo e portanto sabermos a qual esfera pertence uma tessera. Esse parametro será particularmente útil quando for necessário, por motivos que serão discutidos adiante, a remoção de alguma tessera da lista.

Considerando  $k$  superfícies esféricas sobre as quais tenhamos  $n$  triângulos esféricos possuindo as propriedades acima determinadas, a estrutura de dados que representa a tesselação dessa região será definida como:

$$T_\Gamma = \{(C_1, \hat{n}_1, A_{TS1}, 1), \dots, (C_n, \hat{n}_n, A_{TSn}, k)\} \quad (4.10)$$

### 4.1.2 Toro

Um toro  $\tau$  é uma superfície de revolução gerada pela rotação de um circunferência  $\delta$  de raio  $r$  em torno de um eixo coplanar com  $\delta$ . Em particular, caso o eixo de revolução não intercepte a circunferência, a superfície terá a forma de um anel e será chamada de toro de revolução.

Uma parametrização  $\psi(\theta, \phi)$  de  $\tau$  pode ser dada por:

$$\psi = ([R + r\cos\theta]\cos\phi, [R + r\cos\theta]\sin\phi, r\sin\theta) \quad (4.11)$$

onde  $\theta \in [0, 2\pi]$  e  $\phi \in [0, 2\pi]$  e  $R$  é o raio de revolução de uma circunferência de raio  $r$  em torno de um eixo.

Para definir uma tesselação triangular sobre o toro  $\tau$ , definimos uma partição triangular (uma malha) no quadrado  $[0, 2\pi] \times [0, 2\pi]$  do plano  $\theta \times \phi$ . E a partir desta triangulação e utilizando a parametrização do toro, definimos as tesserias triangulares no toro. Para a definição do centro, calculamos o centro dos triangulo no plano  $\theta \times \phi$  e utilizamos a parametrização para calcular o centro da tessera.

O refinamento dessa discretização é obtido simplesmente diminuindo  $n_1$  e  $n_2$ .

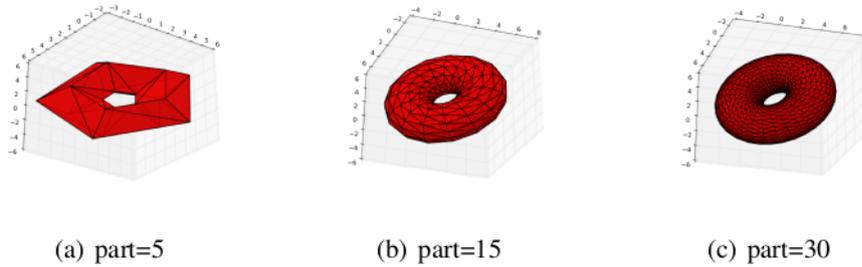


Figura 4.3: Toro de revolução. O parâmetro  $part$  representa o número de partições nas direções  $\theta$  e  $\phi$ .

Uma tesselação de  $\tau$  pode ser determinada considerando os mesmos parâmetros básicos utilizados na esfera: determinar o centro  $C$  de cada região triangular, o vetor normal  $\hat{n}_C$  que passa por  $C$  e a área  $a$  dessa região.

Para o cálculo do centro  $C$ , vamos considerar o ponto médio  $P_m$  da região triangular no plano  $\theta\phi$  e usar  $\psi$  para determinar as coordenadas cartesianas de  $C$ .

$$\theta_C = \frac{\theta_{V1} + \theta_{V2} + \theta_{V3}}{3} \quad (4.12)$$

$$\phi_C = \frac{\phi_{V1} + \phi_{V2} + \phi_{V3}}{3} \quad (4.13)$$

$$P_m = (\theta_{P_m}, \phi_{P_m}) \Rightarrow C = \psi(\theta_{P_m}, \phi_{P_m}) \quad (4.14)$$

Se  $\psi$  a parametrização de  $\tau$  então  $\hat{n} = (-\cos\theta\cos\phi, \sin\phi\cos\theta, -\sin\theta)$ . É suficiente à tesselação calcularmos os vetores que passam pelo centro das tesserias somente.

Para o cálculo da área da região, faz-se necessário o uso de integrais de superfície e novas parametrizações.

Tomando a parametrização  $\psi$  de  $\tau$ , temos:

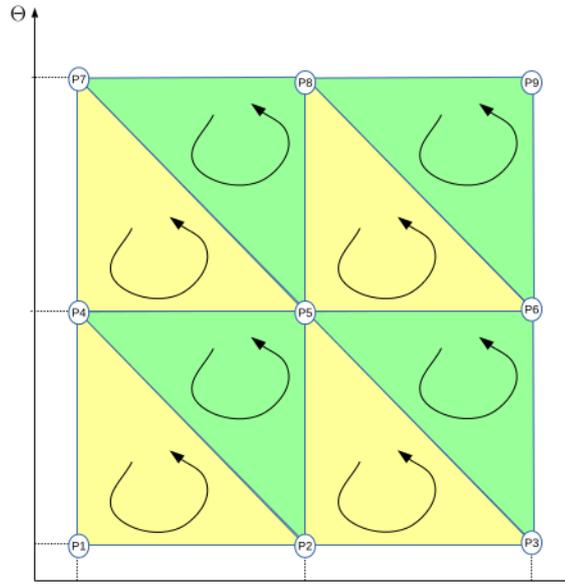


Figura 4.4: Um exemplo de conectividade: Nomea-se os vértices  $P_n$  de coordenadas  $(\theta_n, \phi_n)$ , define-se a forma do elemento e se estabelece como se dá a conexão entre eles, geralmente definindo um sentido preferencial. Usando triângulos e mantendo sempre o sentido anti-horário, uma lista de conectividade válida seria:  $\{(P_1, P_2, P_4), (P_2, P_5, P_4), (P_4, P_5, P_7), (P_5, P_8, P_7), (P_2, P_3, P_5), \dots, (P_6, P_9, P_8)\}$

$$\begin{aligned}\psi_\theta &= (-r \sin\theta \cos\phi, -r \sin\theta \sin\phi, r \cos\theta) \\ \psi_\phi &= (-[R + r \cos\theta] \sin\phi, [R + r \cos\theta] \cos\phi, 0)\end{aligned}\quad (4.15)$$

A área do toro é dada pela integral:

$$\int_0^{2\pi} \int_0^{2\pi} \|r_\theta \times r_\phi\| d\theta d\phi = \int_0^{2\pi} \int_0^{2\pi} (Rr + r^2 \cos\theta) d\theta d\phi = 4\pi^2 Rr \quad (4.16)$$

Para área de cada tessera triangular, será necessário realizar o cálculo em duas etapas. Para efeito de visualização, considere os pontos  $P_1, P_2, P_4, P_5$  da Figura 4.4.

Com esses quatro pontos formamos duas regiões triangulares e precisamos estabelecer a região de integração para ambas. Vamos chamar de *região 1* e *região 2* as regiões de integração delimitada pelos pontos  $P_1, P_2, P_4$  e  $P_2, P_5, P_4$ , respectivamente e a seguir parametrizá-las convenientemente.

Para a *região 1* teremos como integrantes  $\theta_1 \rightarrow \theta_2$  e  $\phi_1 \rightarrow m\theta + n$ , onde  $m = \frac{-(\phi_2 - \phi_1)}{\theta_2 - \theta_1}$  e  $n = -m\theta_2 + \phi_1$ . Calculando a integral sobre essa região e, conseqüentemente a área da tessera definida por ela, encontramos:

$$\begin{aligned}
A_i &= \int_{\theta_1}^{\theta_2} \int_{\phi_1}^{m\theta+n} (Rr + r^2 \cos\theta) d\theta d\phi \\
&= Rrm(\theta_2 - \theta_1) \left( \frac{\theta_1 - \theta_2}{2} \right) + mr^2 [\cos\theta_2 - \cos\theta_1 + \sin\theta_1(\theta_2 - \theta_1)]
\end{aligned} \tag{4.17}$$

Para a *região 2* teremos como integrantes  $\theta_1 \rightarrow \theta_2$  e  $m'\theta + n' \rightarrow \phi_2$ , onde  $m' = \frac{1}{m}$  e  $n' = -m'\theta_1 + \phi_2$ .

Calculando a integral sobre essa região encontramos:

$$A_i = \int_{\theta_1}^{\theta_2} \int_{m'\theta+n'}^{\phi_2} (Rr + r^2 \cos\theta) d\theta d\phi \tag{4.18}$$

$$= -mRr \left[ \frac{-(\theta_2^2 - \theta_1^2)}{2} + \theta_1(\theta_2 - \theta_1) \right] + mr^2 [\cos\theta_1 - \cos\theta_2 + \sin\theta_2(\theta_1 - \theta_2)] \tag{4.19}$$

## 4.2 Patches

Nessa seção serão apresentados os argumentos, definições e parâmetros que permitem a construção, a partir dos elementos geométricos apresentados na seção anterior, dos *patches* utilizados para determinar corretamente a SES.

### 4.2.1 Esferas Covexas

Todos os átomos ou grupos específicos de átomos (CH, por exemplo) serão representados por esferas de raio  $r_i$  obtidos em [2] e centro  $C_i$  obtidos de dados experimentais ou por softwares como Avogadro [14]. Esses parâmetros formam o conjunto de entrada principal do algoritmo.

O primeiro passo para a construção da SES é realizar a tesselação dessas esferas para todos os átomos que compõem a molécula. Será discutido mais especificamente na 4.2.4 que partes ou até mesmo toda a superfície de algumas dessas esferas podem não fazer parte da SES e precisarão ser removidas.

### 4.2.2 Patch de Toro

Considere uma esfera  $\Gamma_{solv}$  de raio  $r_{solv}$ , um par de esferas  $\Gamma_1$  e  $\Gamma_2$  de raios  $r_1$  e  $r_2$ , e  $\alpha$  e  $\beta$  os ângulos entre os centros de  $\Gamma_{solv}$  e  $\Gamma_1$  e  $\Gamma_{solv}$  e  $\Gamma_2$ , respectivamente, em relação a reta que passa pelos centros de  $\Gamma_1$  e  $\Gamma_2$

**Definição 4.2.1.** *Definimos duas esferas  $\Gamma_1$  e  $\Gamma_2$ , referentes aos átomos da molécula, são vizinhas quando uma esfera de solvente  $\Gamma_{solv}$  não puder ocupar inteiramente a região entre  $\Gamma_1$  e*

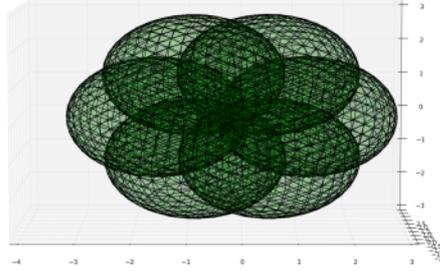


Figura 4.5: Esferas representando os átomos de CH do benzeno

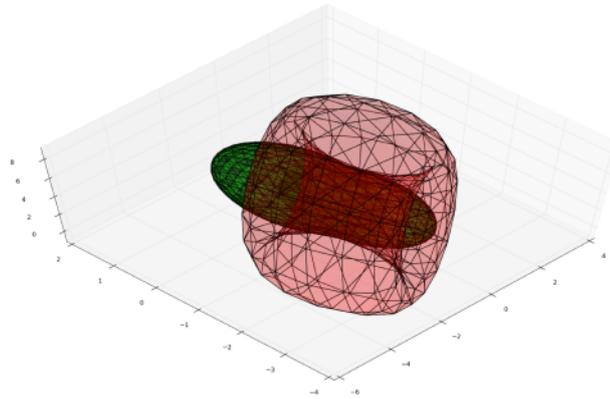


Figura 4.6: Toro sobre uma molécula de CO

$\Gamma_2$ . Ou seja, duas esferas são vizinhas se

$$d(C_1, C_2) < r_1 + r_2 + 2r_{solv} \quad (4.20)$$

Caso  $\Gamma_1$  e  $\Gamma_2$  sejam vizinhos, a trajetória do rolar de  $\Gamma_{solv}$  sobre esse par é descrito por um toro de revolução tal que  $r = r_{solv}$  e  $R = [(r_1 + r_{solv})\sin(\alpha)] = [(r_2 + r_{solv})\sin(\beta)]$

De acordo com a definição da SES, apenas um *patch* ou pedaço desse toro será necessário para sua formação. Em particular, dado um par de esferas vizinhas e uma discretização do plano  $\theta \times \phi$  com  $ndiv = N$ , definimos o *patch*  $T_{ij}$  limitando o domínio discreto em  $\theta$  tal que

$$Dom(\theta) = \{\theta_0, \theta_0 + \Delta\theta, \dots, \theta_0 + (n - 1)\Delta\theta\} \quad (4.21)$$

onde  $\theta_0 = \beta + \frac{\pi}{2}$ ,  $\theta_N = \theta_0 + \pi - (\alpha + \beta)$ ,  $\Delta\theta = \frac{(\theta_N - \theta_0)}{N-1}$  e o domínio discreto em  $\phi$  tal que

$$Dom(\phi) = \{\phi_0, \phi_0 + \Delta\theta, \dots, \phi_0 + (n-1)\Delta\theta, \phi_N\} \quad (4.22)$$

onde  $\phi_0 = 0$ ,  $\phi_N = 2\pi$  e  $\Delta\theta = \frac{2\pi}{N}$ .

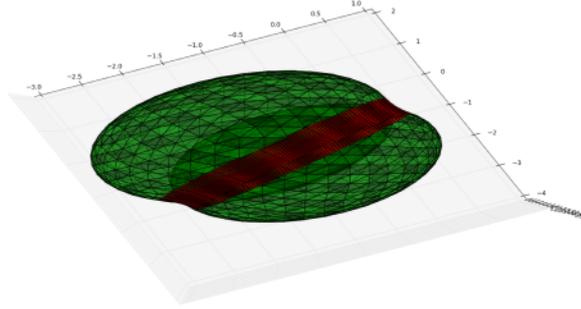


Figura 4.7: Patch de toro sobre uma molécula de CO

### 4.2.3 Patch de Esfera concava

Ao rolar a esfera de solvente em volta do soluto, poderá ocorrer de a esfera de solvente tangenciar o soluto em 3 ou mais átomos simultaneamente. Isto define uma superfície esférica concava.

Como bastam tres esferas para definir o centro desta esfera de solvente, considere um trio de esferas  $\Gamma_1(r_1, C_1)$ ,  $\Gamma_2(r_2, C_2)$ ,  $\Gamma_3(r_3, C_3)$  vizinhas entre si, ou seja, vizinhas dois a dois.

Seja  $\Gamma_{solv}$  uma esfera de raio  $r_{solv}$  tangente simultaneamente a  $\Gamma_1$ ,  $\Gamma_2$  e  $\Gamma_3$  em  $P_1$ ,  $P_2$  e  $P_3$  respectivamente. Caso exista, o centro  $C_{solv}$  de  $\Gamma_{solv}$  será solução do sistema[29]

$$\begin{cases} d(C_1, C_{solv})^2 = (r_1 + r_{solv})^2 \\ d(C_2, C_{solv})^2 = (r_2 + r_{solv})^2 \\ d(C_3, C_{solv})^2 = (r_3 + r_{solv})^2 \end{cases} \quad (4.23)$$

Os pontos  $P_1$ ,  $P_2$  e  $P_3$  serão obviamente  $\Gamma_1 \cap \Gamma_{solv}$ ,  $\Gamma_2 \cap \Gamma_{solv}$  e  $\Gamma_3 \cap \Gamma_{solv}$  respectivamente. Assim, temos:

$$P_1 = C_1 + \frac{C_{solv}C_1}{|C_{solv}C_1|}r_1 \quad (4.24)$$

$$P_2 = C_2 + \frac{C_{solv}C_2}{|C_{solv}C_2|}r_2 \quad (4.25)$$

$$P_3 = C_3 + \frac{C_{solv}C_3}{|C_{solv}C_1|}r_3 \quad (4.26)$$

Esses pontos definem um plano  $\alpha$  a partir do qual será dada uma aproximação da superfície necessária para limitar  $\Gamma_{solv}$  e formar o patch esférico que conecta essas três esferas vizinhas.

#### 4.2.4 Mecanismos de remoção

A partir das esferas centradas nos átomos e dos patches conseguimos, com boa aproximação, cobrir a região da SES. Entretanto algumas das tesseras formadas podem pertencer a dois objetos geométricos simultaneamente ou podem não pertencer mais a SES, mas sim ao seu interior. Por esse motivo faz-se necessário um mecanismo para remover essas tesseras excedentes.

Seja  $P$  um ponto sobre uma esfera convexa  $S$ . Define-se para todo  $P \in S$  um centro  $C_s$  de uma esfera  $\xi$  de raio  $r_{solv}$  tangente a  $S$  em  $P$ . Testa-se se um ponto  $T$  sobre outro *patch* intercepta  $\xi$ . Caso afirmativo, remove-se a tessera que contém  $T$ . O processo é repertido para todas as esferas convexas obtidas na fase inicial da tesselação.

#### 4.2.5 Implementação

O algoritmo proposto para tesselação foi implementado em Python. Além das funções *built-in*, foram usados os pacotes *matplotlib* [16] para a parte gráfica e *numpy* [31] para as estruturas e cálculos relacionados à álgebra linear. O código-fonte pode ser encontrado em [https://drive.google.com/open?id=0B\\_21i1LdX5aVMVBINzRhTWp0M1k](https://drive.google.com/open?id=0B_21i1LdX5aVMVBINzRhTWp0M1k).

## CAPÍTULO 5

---

### Resultados Numéricos

---

Para tornar possível uma avaliação para o MultiPatch foi necessária a utilização do software GAMESS [27] que possui implementado o modelo PCM [22], com algumas de suas variações. Além de possuir a implementação de algumas versões do GEPOL. A partir desse software obtivemos um conjunto de moléculas solvatadas cujo resultado previsto teoricamente se mostrou acurado em comparação com os dados experimentais. Esse fato possibilitou a formação de um grupo de controle. Por ser um software de código aberto foi possível também alterar o código-fonte e inserir a nova tesselação e desse modo comparar os resultados.

#### 5.1 GAMESS

O GAMESS (abreviação de General Atomic and Molecular Electronic Structure System) é um pacote geral para cálculos quanto-mecânicos. Atualmente é mantido pelos membros do grupo de pesquisa Gordon na Universidade do Estado de Iowa (<http://www.msg.ameslab.gov/games/index.html>).

Trata-se de um software de código aberto derivado principalmente do HONDO 5 e que permite um grande número de cálculos resultantes da quanto-mecânicos (RHF, UHF, ROHF, GVB, MCSCF), sendo bastante conhecido pela comunidade científica da área de solvatação contínua. Possui alguns métodos para se obter a tesselação, todas baseadas no GEPOL [25, 30, 20].

A tesselação obtida pelo MultiPatch é testada com o uso desse software. Para tal, nosso código gera a tesselação e salva os dados necessários em um arquivo binário. Este arquivo substitui, durante a execução do GAMESS, aquele utilizado pelo código original. Foi necessário editar a rotina principal de geração de malha do GAMESS, onde implementamos a leitura dos dados do arquivo gerado. Tal implementação de leitura foi testada utilizando a

própria tesselação gerada pelo GAMESS.

## 5.2 Análise de dados

### 5.2.1 Testes com GEPOL

Como experimento numérico base, testamos o comportamento do modelo e do software utilizando solvente e soluto compostos pelas mesmas substâncias.

Na tabela a seguir estão listados os resultados obtidos para o cálculo da energia livre de solvatação desse experimento numérico para um conjunto de moléculas com características diferentes, além da comparação desse resultado com os dados experimentais para a energia livre de solvatação [21]. Para a formação desse conjunto de moléculas respeitaram-se duas características: uso de espécies químicas diferentes e a sua prévia implementação no GAMESS (Apêndice).

| MOLÉCULA                | $\Delta G_{solu}$ -NUMÉRICO(kJ/mol) | $\Delta G_{solu}$ -EXPERIMENTAL(kJ/mol) | $\epsilon$ |
|-------------------------|-------------------------------------|---|------------|
| AGUA                    | -26.46                              | -26.50                                  | 78.39      |
| DMSO                    | -32.81                              | -32.02                                  | 46.7       |
| METANOL                 | -20.33                              | -17.82                                  | 32.63      |
| ETANOL                  | -21.25                              | -16.88                                  | 24.55      |
| ANILINA                 | 11.76                               | -31.27                                  | 6.89       |
| CLOROFÓRMIO             | 23.20                               | -17.51                                  | 4.90       |
| CICLOHEXANO             | 4.01                                | -18.51                                  | 2.02       |
| TOLUENO                 | 2.74                                | -21.61                                  | 2.38       |
| BENZENO                 | -6.82                               | -19.06                                  | 2.25       |
| TETRACLORETO DE CARBONO | 38.41                               | -18.41                                  | 2.23       |
| NHEPTANO                | -4.38                               | -19.64                                  | 1.92       |

Tabela 5.1: Grupo de Controle e energia de solvatação.

Na Tabela 5.1 é possível notar que o modelo funciona de maneira acurada quando se tem uma grande influência da componente eletrostática, isto é, quando o soluto é uma espécie polar e o solvente possui grande valor de constante dielétrica  $\epsilon$ . As quatro primeiras moléculas da Tabela 5.1 foram escolhidas para formar o grupo de controle: Água, Dimetilsulfóxido(DMSO), Metanol e Etanol. Uma vez que para esse grupo já existe uma boa correlação entre o cálculo numérico e a medida experimental, como pode ser visto na Figura 5.1, espera-se que o método proposto por esse trabalho se comporte tão bem quanto ou melhor que o GEPOL nesses casos antes de ser submetido aos casos onde o GEPOL apresenta dificuldades.

### 5.2.2 Testes com MultiPatch

A seguir encontram-se os resultados obtidos na execução do GAMESS usando as equações que definem os métodos DPCM, IEF-PCM e COSMO vistos no capítulo 2 e os métodos de tesselação GEPOL-RT, GEPOL-AS, GEPOL-GB, FIXPVA ([25, 20, 30]) e o MultiPatch. Uma observação importante: com a tesselação gerada pelo método MultiPatch foi calculada apenas a componente eletrostática  $\Delta G_{el}^*$ . Para compor a energia livre de solvatação  $\Delta G_s^*$  foram

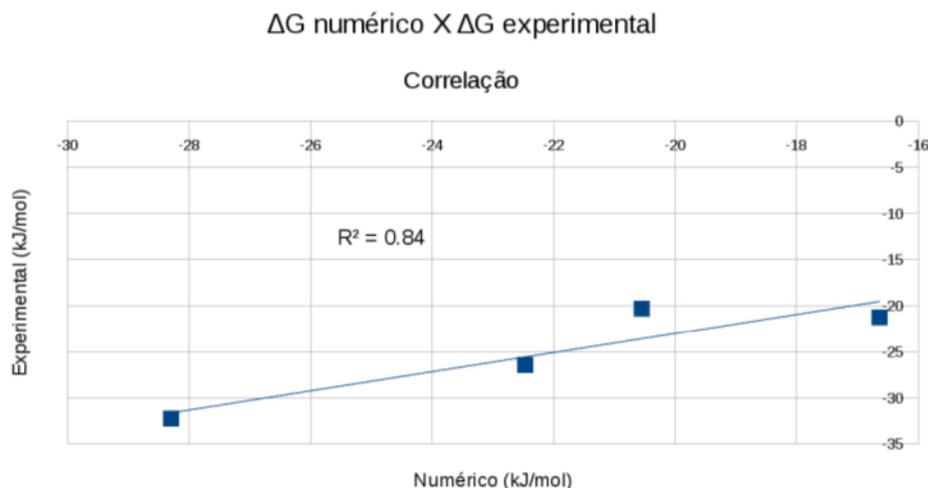


Figura 5.1: Correlação da energia livre de solvatação experimental e numérica para as moléculas do grupo de controle: DMSO, Água, Metanol, Etanol.

utilizados os valores obtidos para as componentes de solvatação, dispersão e repulsão obtidos utilizando a tesselação gerada pelo FIXPVA. Além disso, um problema ainda não solucionado na implementação do patch convexo limitou a geração da cavidade via MultiPatch aos patches concavos e tóricos.

Na tabela 5.2 encontram-se os testes realizados utilizando o método DPCM, onde vê-se claramente como os resultados são afetados pela forma da cavidade. Usando as cavidades geradas pelos métodos FIXPVA e MultiPatch, a energia livre de solvatação  $\Delta G_s^*$  apresentou valores positivos e grandes em todos os casos. Fisicamente esse resultado representa a existência de uma forte repulsão entre as moléculas em todas as soluções, o que obviamente é um absurdo mesmo sendo os dados experimentais apresentados.

Na tabela 5.3 encontram-se os testes realizados utilizando o método IEFPCM. Apesar de obtido a partir do DPCM, esse método apresenta melhor flexibilidade quanto a forma da cavidade como pode ser observado. Apesar de um resultado não acurado no teste com o Etanol, o valor da energia obtido com o MultiPatch para a Água apresentou menor erro e mostrou-se compatível aos demais nos testes com o Metanol e o DMSO.

Na tabela 5.4 encontram-se os teste realizados utilizando o método COSMO. Como já apresentado, trata-se de uma aproximação do método IEFPCM e portanto fazia sentido esperar que apresentasse também boa flexibilidade quanto a forma da cavidade. Um fato notável nos resultado apresentados nessa tabela pode ser observado a partir de uma comparação direta dos valores de energia obtidos utilizando a cavidade gerada pelo FIXPVA, projetado para métodos desse tipo (C-PCM especificamente) [30] e os obtidos com o MultiPatch (Figura 5.2).

| ÁGUA EM ÁGUA       | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
|--------------------|-------------------------|-------------------|------------------------------|-------|-------|
| EXPERIMENTAL       | -26.46                  | -                 | -                            | -     | 78.39 |
| FIXPVA             | 199566.21               | 754318.48         | 41.27319012                  | 87    | 78.39 |
| GEPOL-GB           | -22.01                  | 16.81             | 48.08801955                  | 120   | 78.39 |
| GEPOL-RT           | -21.81                  | 17.57             | 46.38991164                  | 3296  | 78.39 |
| GEPOL-AS           | -22.01                  | 16.81             | 48.08801955                  | 120   | 78.39 |
| MultiPatch         | 199569.34               | 754330.33         | 47.17455982                  | 465   | 78.39 |
| METANOL EM METANOL | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
| EXPERIMENTAL       | -21.25                  | -                 | -                            | -     | 24.55 |
| FIXPVA             | 302055.13               | 1421535.89        | 50.74842282                  | 134   | 24.55 |
| GEPOL-GB           | -20.69                  | 2.62              | 74.26649275                  | 213   | 24.55 |
| GEPOL-RT           | -20.42                  | 3.90              | 70.18941985                  | 7234  | 24.55 |
| GEPOL-AS           | -20.69                  | 2.62              | 74.26649275                  | 213   | 24.55 |
| MultiPatch         | -1492711.85             | -7024426.34       | 63.98074308                  | 630   | 24.55 |
| ETANOL EM ETANOL   | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
| EXPERIMENTAL       | -20.33                  | -                 | -                            | -     | 32.63 |
| FIXPVA             | 404563.29               | 1990081.73        | 62.21778605                  | 174   | 32.63 |
| GEPOL-GB           | -19.11                  | 6.01              | 97.86844620                  | 295   | 32.63 |
| GEPOL-RT           | -18.74                  | 7.80              | 91.54233816                  | 9533  | 32.63 |
| GEPOL-AS           | -18,32                  | 9.90              | 96.60022775                  | 295   | 32.63 |
| MultiPatch         | -4.41                   | 78.29             | 72.43421387                  | 684   | 32.63 |
| DMSO EM DMSO       | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
| EXPERIMENTAL       | -32.28                  | -                 | -                            | -     | 46.70 |
| FIXPVA             | 1448087.08              | 4486119.45        | 89.42775658                  | 198   | 46.70 |
| GEPOL-GB           | -31.32                  | 2.98              | 122.32561386                 | 333   | 46.70 |
| GEPOL-RT           | -30.76                  | 4.71              | 115.58080572                 | 10483 | 46.70 |
| GEPOL-AS           | -30.74                  | 4.77              | 120.80458731                 | 333   | 46.70 |
| MultiPatch         | -5818022.70             | -18023514.32      | 92.91894486                  | 793   | 46.70 |

Tabela 5.2: Resultados utilizando o método DPCM não-iterativo

| ÁGUA EM ÁGUA       | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
|--------------------|-------------------------|-------------------|------------------------------|-------|-------|
| EXPERIMENTAL       | -26.46                  | -                 | -                            | -     | 78,39 |
| FIXPVA             | -23.36                  | 11.70             | 41.27319012                  | 87    | 78,39 |
| GEPOL-GB           | -22.74                  | 14.06             | 48.08801955                  | 120   | 78,39 |
| GEPOL-RT           | -22.87                  | 13.58             | 46.38991164                  | 3296  | 78,39 |
| GEPOL-AS           | -22.74                  | 14.06             | 48.08801955                  | 120   | 78,39 |
| MultiPatch         | -23.78                  | 10.12             | 47.17455982                  | 465   | 78,39 |
| METANOL EM METANOL | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
| EXPERIMENTAL       | -21.25                  | -                 | -                            | -     | 24,55 |
| FIXPVA             | -20.56                  | 3.27              | 50.74842282                  | 134   | 24,55 |
| GEPOL-GB           | -20.62                  | 2.97              | 74.26649275                  | 213   | 24,55 |
| GEPOL-RT           | -20.74                  | 2.42              | 70.18941985                  | 7234  | 24,55 |
| GEPOL-AS           | -20.62                  | 2.97              | 74.26649275                  | 213   | 24,55 |
| MultiPatch         | -20.55                  | 3.29              | 63.98074308                  | 630   | 24,55 |
| ETANOL EM ETANOL   | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
| EXPERIMENTAL       | -20.33                  | -                 | -                            | -     | 32,63 |
| FIXPVA             | -16.63                  | 18.19             | 62.21778605                  | 174   | 32,63 |
| GEPOL-GB           | -18.59                  | 8.56              | 97.86844620                  | 295   | 32,63 |
| GEPOL-RT           | -18.78                  | 7.61              | 91.54233816                  | 9533  | 32,63 |
| GEPOL-AS           | -18.48                  | 9,08              | 96.60022775                  | 295   | 32,63 |
| MultiPatch         | -9.84                   | 51.57             | 72.43421387                  | 684   | 32,63 |
| DMSO EM DMSO       | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
| EXPERIMENTAL       | -32.28                  | -                 | -                            | -     | 46.70 |
| FIXPVA             | -29.70                  | 7.99              | 89.42775658                  | 198   | 46.70 |
| GEPOL-GB           | -30.50                  | 5.52              | 122.32561386                 | 333   | 46.70 |
| GEPOL-RT           | -30.70                  | 4.89              | 115.58080572                 | 10483 | 46.70 |
| GEPOL-AS           | -30.23                  | 6.35              | 120.80458731                 | 333   | 46.70 |
| MultiPatch         | -28.69                  | 11.11             | 92.91894486                  | 793   | 46.70 |

Tabela 5.3: Resultados utilizando o método IEF-PCM não-iterativo

| ÁGUA EM ÁGUA       | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
|--------------------|-------------------------|-------------------|------------------------------|-------|-------|
| EXPERIMENTAL       | -26.46                  | -                 | -                            | -     | 78.39 |
| FIXPVA             | -23.53                  | 11.09             | 41.27319012                  | 87    | 78.39 |
| GEPOL-GB           | -22.91                  | 13.43             | 48.08801955                  | 120   | 78.39 |
| GEPOL-RT           | -23.03                  | 12.95             | 46.38991164                  | 3296  | 78.39 |
| GEPOL-AS           | -22.91                  | 13.43             | 48.08801955                  | 120   | 78.39 |
| MultiPatch         | -24.04                  | 9.16              | 47.17455982                  | 465   | 78.39 |
| METANOL EM METANOL | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
| EXPERIMENTAL       | -21.25                  | -                 | -                            | -     | 24,55 |
| FIXPVA             | -20.83                  | 1.97              | 50.74842282                  | 134   | 24,55 |
| GEPOL-GB           | -20.92                  | 1.57              | 74.26649275                  | 213   | 24,55 |
| GEPOL-RT           | -21.04                  | 0.98              | 70.18941985                  | 7234  | 24,55 |
| GEPOL-AS           | -20.92                  | 1.57              | 74.26649275                  | 213   | 24,55 |
| MultiPatch         | -21.02                  | 1.08              | 63.98074308                  | 630   | 24,55 |
| ETANOL EM ETANOL   | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
| EXPERIMENTAL       | -20.33                  | -                 | -                            | -     | 32.63 |
| FIXPVA             | -16.99                  | 16.42             | 62.21778605                  | 174   | 32.63 |
| GEPOL-GB           | -19.05                  | 6.32              | 97.86844620                  | 295   | 32.63 |
| GEPOL-RT           | -19.21                  | 5.49              | 91.54233816                  | 9533  | 32.63 |
| GEPOL-AS           | -18.90                  | 7.02              | 96.60022775                  | 295   | 32.63 |
| MultiPatch         | -17.07                  | 16.01             | 72.43421387                  | 684   | 32.63 |
| DMSO EM DMSO       | $\Delta G_s^*$ (kJ/mol) | Erro Relativo (%) | Área Total (Å <sup>2</sup> ) | NTS   | €     |
| EXPERIMENTAL       | -32.28                  | -                 | -                            | -     | 46.70 |
| FIXPVA             | -29.70                  | 7.99              | 89.42775658                  | 198   | 46.70 |
| GEPOL-GB           | -30.88                  | 4.34              | 122.32561386                 | 333   | 46.70 |
| GEPOL-RT           | -31.12                  | 3.60              | 115.58080572                 | 10483 | 46.70 |
| GEPOL-AS           | -30.61                  | 5.16              | 120.80458731                 | 333   | 46.70 |
| MultiPatch         | -29.46                  | 8.74              | 92.91894486                  | 793   | 46.70 |

Tabela 5.4: Resultados utilizando o método COSMO não-iterativo

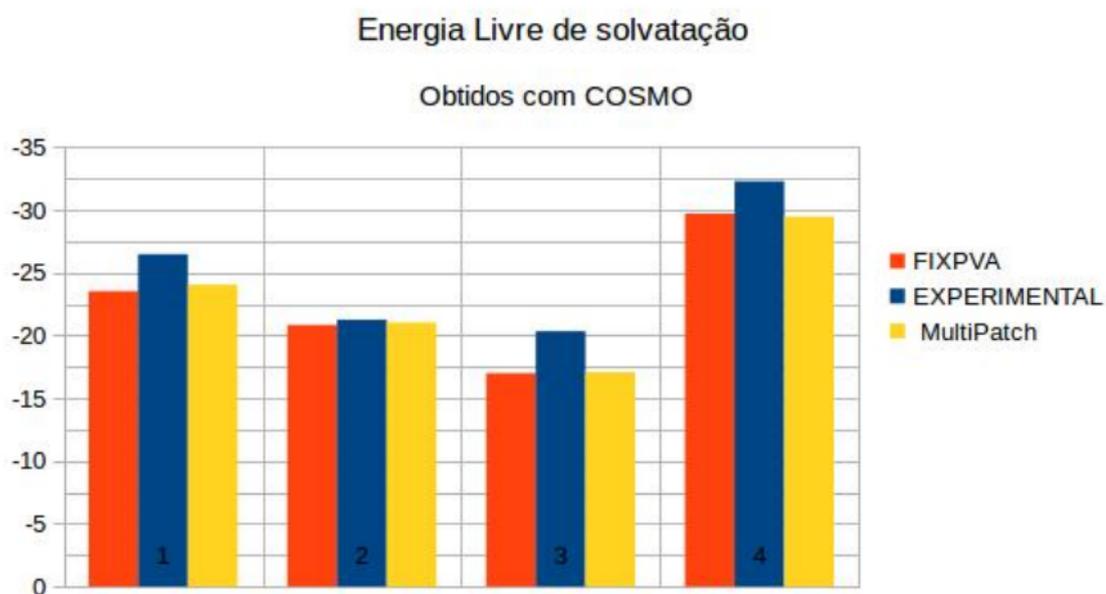


Figura 5.2: Comparação na energia livre de solvatação  $\Delta G_s^*$  obtida a partir das cavidades geradas com o FIXPVA e o MultiPatch usando o método COSMO.

## CAPÍTULO 6

---

### Conclusões

---

No Capítulo 3 estabelecemos as características básicas para a nova metodologia.

Um resultado parcial foi obtido quanto a continuidade da tesselação, pois ainda não se implementou um refinamento apropriado na região onde ocorre a união dos patches. Entretanto a quantidade de ajustes desse tipo é bem inferior ao necessário pelo GEPOL.

Apesar de não apresentarem áreas iguais, todas as tesseras são triangulares como esperado. Algumas soluções para melhoria do *ratio aspect* das tesseras foram discutidas, mas ainda não foram implementadas.

Nenhum experimento relacionado a conformação ou otimização da geometria molecular não foram realizados. Entretanto, os componentes de cada tessera, ou seja, sua área, centro e normal foram obtidos de forma analítica, favorecendo o cálculo de derivadas necessárias para o processo de otimização.

O software MultiPatch apresenta rápida execução mesmo sendo desenvolvido a partir de uma linguagem interpretada. Existe um interesse quanto a tradução o mesmo em  $C++$ , o tornando ainda mais eficiente.

Quanto aos resultados numéricos, conclui-se que o MultiPatch é uma alternativa viável geração da SES como apresentada na definição uma vez que apresenta resultados satisfatórios quando comparados aos métodos implementados no GAMESS. Quando utilizado com o método COSMO, apresentou resultados mais acurado ou equivalentes ao FIXPXA para cálculo da componente eletrostática. Essa comparação é relevante pois esse métodos é atualmente o gerador de malha padrão no GAMESS e foi desenvolvido tendo em vista esse método de solução.

Esperamos dar continuidade ao desenvolvimento de desse software e, uma vez na versão final, contribuir para o desenvolvimento da área.

# APÊNDICE A

## Entrada e saída do GAMESS

Nesse apêndice encontram-se algumas entradas e saídas relativas a simulação no GAMESS das moléculas que formaram o grupo de controle para avaliação dos resultados do novo método proposto. Em todos os exemplos, o modelo PCM utilizado, definido pela opção IEF = -3, trata o solvente como um dielétrico isotrópico e resolve os sistemas relacionados a esse modelo através de métodos iterativos. Por limitação do software, todos foram calculado sem otimização de geometria (*single point*, opção RUNTYP =ENERGY) na fase líquida<sup>1</sup>, com soluto e solvente formado pela mesma espécie química e sendo habilitado o cálculo para as componentes de dispersão/repulsão (IDISP) e de cavitação (ICAV).

Como a apresentação das saídas na sua totalidade se tornaria inviável, coube enfatizar apenas a parcela da saída correspondente ao cálculo da energia livre de solvatação  $\Delta G_s^*$ .

### Exemplo A.0.1.<sup>2</sup>

```

1  agua.inp
2  !File created by the GAMESS Input Deck Generator Plugin for Avogadro
3  $BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 NPFUNC=1 $END
4  $PCM SOLVNT=WATER IEF=-3 IDISP=1 ICAV=1 $END
5  $CONTRL SCFTYP=RHF RUNTYP=ENERGY$END
6  $DATA
7  Title
8  Cl
9  H      1.0      0.61665      -0.02647      0.47106
10 O      8.0      0.14969      0.03034      -0.34628
11 H      1.0      -0.76633      -0.00388      -0.12478

```

<sup>1</sup>Em uma etapa anterior, foi realizada a otimização da geometria em fase gasosa.

<sup>2</sup>Conforme consta no arquivo de entrada, a geometria inicial foi obtida pelo software Avogadro [14]

```

12  $END
13
14
15  Arquivo de saída
16
17  agua.log
18
19  -----
20  ----- RESULTS OF PCM CALCULATION -----
21  -----
22
23  FREE ENERGY IN SOLVENT = <PSI| H(0)+V/2 |PSI>      =
      -76.0333317672 A.U.
24  INTERNAL ENERGY IN SOLVENT = <PSI| H(0) |PSI>      =
      -76.0227391041 A.U.
25  DELTA INTERNAL ENERGY = <D-PSI| H(0) |D-PSI>      =
      0.0000000000 A.U.
26  ELECTROSTATIC INTERACTION =                          =
      -0.0105926631 A.U.
27  PIEROTTI CAVITATION ENERGY =                        =
      0.0080923832 A.U.
28  DISPERSION FREE ENERGY =                            =
      -0.0087022095 A.U.
29  REPULSION FREE ENERGY =                             =
      0.0017684793 A.U.
30  TOTAL INTERACTION (DELTA + ES + CAV + DISP + REP) =
      -0.0094340101 A.U.
31  TOTAL FREE ENERGY IN SOLVENT =                      =
      -76.0321731142 A.U.
32
33  FREE ENERGY IN SOLVENT      =      -47711.64 KCAL/MOL
34  INTERNAL ENERGY IN SOLVENT  =      -47704.99 KCAL/MOL
35  DELTA INTERNAL ENERGY       =           0.00 KCAL/MOL
36  ELECTROSTATIC INTERACTION     =      -6.65 KCAL/MOL
37  PIEROTTI CAVITATION ENERGY   =       5.08 KCAL/MOL
38  DISPERSION FREE ENERGY       =      -5.46 KCAL/MOL
39  REPULSION FREE ENERGY        =       1.11 KCAL/MOL
40  TOTAL INTERACTION             =      -5.92 KCAL/MOL
41  TOTAL FREE ENERGY IN SOLVENT =     -47710.91 KCAL/MOL
42
43  ----- ENERGY CHANGE FROM GAS PHASE TO SOLVENT -----

```

---

|    |             |                 |                         |
|----|-------------|-----------------|-------------------------|
| 44 | <i>ELEC</i> | <i>ELEC+CAV</i> | <i>ELEC+CAV+DIS+REP</i> |
| 45 | -6.097      | -1.019          | -5.370 <i>KCAL/MOL</i>  |

---

## Referências Bibliográficas

---

- [1] A. BEN-NAIM, *Solvation Thermodynamics*, Springer US, Jerusalem, Israel, 1 ed., 1987.
- [2] A. BONDI, *van der waals volumes and radii*, J. Phys. Chem., 68 (1964), pp. 441–451.
- [3] D. S. BRUCE COHEN, *The gamma function, factorials and the volumes of n-balls-*, 2001.
- [4] J. D. C. A. BREBBIA, *Boundary elements: an introductory course*, Computational Mechanics, 2 ed., 1992.
- [5] H. B. CALLEN, *Thermodynamics and an Introduction To Thermostatistics*, Wiley, Pennsylvania, 2 ed., 1985.
- [6] E. CANCÈS, B. MENNUCCI, AND J. TOMASI, *A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics*, Journal of Computational Chemistry, 107 (1997), pp. 3032–1776.
- [7] J. T. CHRISTIAN SILVIO POMELLI, *Defpol: New procedure to build molecular surfaces and its use in continuum solvation methods*, Journal of Computational Chemistry, 19 (1998), pp. 1758–1776.
- [8] B. C.J.F., *Theory of Electric Polarization, Vol. 1: Dielectrics in Static Fields*, Elsevier, Amsterdam, Holanda, 2 ed., 1973.
- [9] H. S. M. COEXTER, *Regular Complex Polytopes*, Cambridge U. P., Cambridge, 1974.
- [10] M. L. CONNOLLY, *Analytical molecular surface calculation*, Journal of Applied Crystallography, 16 (1983), pp. 548–558.
- [11] C. J. CRAMER, *Essentials of computational chemistry : theories and models*, John Wiley & Sons, Chichester, West Sussex, England, Hoboken, NJ, 2004.

- [12] L. C. EVANS, *Partial Differential Equations: Second Edition*, American Mathematical Society, Providence, Rhode Island, 1998.
- [13] W. HACKBUSCH, *Integral Equations: Theory and Numerical Treatment*, Birkhäuser Basel, 1 ed., 1995.
- [14] M. D. HANWELL, D. E. CURTIS, D. C. LONIE, T. VANDERMEERSCH, E. ZUREK, AND G. R. HUTCHISON, *Avogadro: an advanced semantic chemical editor, visualization, and analysis platform*, Journal of Cheminformatics, 4 (2012), pp. 1–17.
- [15] T. J. HUGHES, *The finite element method Linear static and dynamic finite element analysis*, Prentice-Hall International, 1987.
- [16] J. D. HUNTER, *Matplotlib: A 2d graphics environment*, Computing In Science & Engineering, 9 (2007), pp. 90–95.
- [17] R. C. JACOPO TOMASI, BENEDETTA MENNUCCI, *Quantum mechanical continuum models*, Chemical Reviews, 105 (2005), pp. 2999–3094.
- [18] O. D. KELLOGG, *Foundations of potential theory*, Verlag Julius Springer, 1929.
- [19] A. KLAMT AND G. SCHÜRSMANN, *Cosmo: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient*, J. Chem. Soc., Perkin Trans. 2, (1993), pp. 799–805.
- [20] H. LI AND J. H. JENSEN, *Improving the efficiency and convergence of geometry optimization with the polarizable continuum model: New energy gradients and molecular surface tessellation*, Journal of Computational Chemistry, 25 (2004), pp. 1449–1462.
- [21] Y. MARCUS, *The Properties of Solvents*, vol. Volume 4 of Wiley Series in Solutions Chemistry, Wiley, 1998.
- [22] B. MENNUCCI AND R. CAMMI, eds., *Continuum Solvation Models in Chemical Physics*, John Wiley & Sons, 2007, ch. Modern Theories of Continuum Models, pp. 1–123.
- [23] H. NUSSENZVEIG, *Curso de Física: v. 3-Eletromagnetismo*, Edgard Blücher, São Paulo, 2001.
- [24] D. OF CHEMICAL NOMENCLATURE, S. R. I. U. OF PURE, A. CHEMISTRY, M. NIC, J. JIRAT, AND B. KOSATA, *IUPAC goldbook*, IUPAC, 2006.
- [25] J.-L. PASCUAL-AHUIR, E. SILLA, AND I. TUÑÓN, *Gepol: An improved description of molecular surfaces. iii. a new algorithm for the computation of a solvent- excluding surface*, Journal of Computational Chemistry, 15 (1994), pp. 1127–1139.

- [26] C. REICHARDT, *Solvents and Solvent Effects in Organic Chemistry*, John Wiley and Sons, Weinheim, Alemanha, 3 ed., 2003.
- [27] M. W. SCHMIDT, K. K. BALDRIDGE, J. A. BOATZ, S. T. ELBERT, M. S. GORDON, J. H. JENSEN, S. KOSEKI, N. MATSUNAGA, K. A. NGUYEN, S. SU, T. L. WINDUS, M. DUPUIS, AND J. A. MONTGOMERY, *General atomic and molecular electronic structure system*, *Journal of Computational Chemistry*, 14 (1993), pp. 1347–1363.
- [28] J. STEWART, *Calculus*, Cengage Learning, 2011.
- [29] B. N. STURGESS AND F. T. CAREY, *Trilateration*, Springer US, Boston, MA, 1987, pp. 340–389.
- [30] P. SU AND H. LI, *Continuous and smooth potential energy surface for conductorlike screening solvation model using fixed points with variable area*, *The Journal of Chemical Physics*, 130 (2009).
- [31] S. V. D. WALT, S. C. COLBERT, AND G. VAROQUAUX, *The numpy array: A structure for efficient numerical computation*, *Computing in Science and Engg.*, 13 (2011), pp. 22–30.
- [32] P. I. WOLFGANG HACKBUSCH, R. FADIMAN, *Elliptic differential equations: Theory and numerical treatment*, Springer Series in Computational Mathematics, Springer, 2003.