Numerical Computing of Molecular Electrostatics through Boundary Integral Equations

October 10, 1996

Jie Liang[1] and Shankar Subramaniam[2]

Summary. In continuum approaches to molecular electrostatics, boundary element method (BEM) can provide accurate solutions to the Poisson-Boltzmann equation. However, the numerical aspects of this method are fraught with difficulties. We describe our approach applying an alpha shape-based method to generate a high quality mesh, which represents the shape and topology of the molecule precisely. We also describe the analytical method to map points from the planar mesh to their exact locations on the surface of the molecule. We demonstrate that derivative boundary integral formulation has numerical advantages over the nonderivative formulation: the well-conditioned influence matrix can be maintained without deteriorating condition number when the number of the mesh elements scale up. Singular integrand kernels are characteristics of the BEM. Their accurate integration is an important issue. We describe variable transformations that allow accurate numerical integraton, the only plausible integral

[1] National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. (217)244–5818, fax (217)244–2909, jliang@ncsa.uiuc.edu

[2] National Center for Supercomputing Applications, Beckman Institute for Advanced Science and Technology, Department of Molecular & Integrative Physiology, Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. (217)244–4489, fax (217)244–2909, shankar@ncsa.uiuc.edu

evaluation method when using curve-shaped boundary elements.

Keywords. molecular shape representation, continuum model, alpha shape, Poisson-Boltzmann equation

1 Introduction

Electrostatic interactions play an important role in macromolecular structure and funtion. Continuum approach provides a convenient route for understanding the molecular electrostatic interactions [1, 2, 3, 4, 5]. It models a molecule as a cavity of molecular shape of low interior dielectric constant embedded in a continuous medium of solvent of high dielectric constant. The interior of the cavity contains point charges representing charge distrubtions on atoms in the molecule. Poisson equation or Poisson-Boltzmann equaiton are then solved for the elcetrostatic potentials over the space. Two widely used methods for solving the continuum electrostatic problems are *the finite difference method* (FDM) and *the boundary element method* (BEM). FDM employs a box of three dimensional cubic grids, where the solute molecule and a portion of surrounding continuum solvent are contained. The electrostatic potentials are then directly solved from the partial differential equation on the grid points [6, 1, 7, 4, 5]. BEM uses Green's second identity to transform the elliptic partial differential equation over a molecule-containing volume to integral equations over the surface of the molecule [8, 9, 10, 11, 12, 13, 14, 15]. With a discretized surface description, the integral equations are then solved numerically. BEM can also be thought of as computing a distribution of the induced point charges (single layer) due to the polarization of the solvent and a distribution of the induced point dipoles (double layer) due to the effects of the ions present in the solvent.

Boundary element method can treat both the molecular shape and the interior charges very accurately. It discretizes the surface of the molecule and relies only on the surface integrals of the molecule. The collection of the discrete surface elements allows the possibility to represent the shape of the molecule precisely. This is in contrast to the nature of the molecular boundary imposed by cubic volume elements in a finite difference method. The interior charges can also be placed at their exact locations, since these charges no longer need to coincide with grid points. Therefore, no distortion in molecular charge distribution is introduced, unlike in FDM, where artificial distribution of fractional charges onto neighbouring grid points is used. As a result, BEM has no need for special remedies such as focusing and rotational averaging, which are commonly practiced in finite difference methods. The self energy artefacts [16] in calculation of reaction potential also do not exist in BEM. In addition, the space outside the molecule does not have to be discretized in BEM, and therefore the potential at places distant from

the molecule can be computed accurately. Furthermore, since there is no finite volume box involved, there are no artificial boudary conditions as imposed in finite difference method when modeling systems infinite relative to molecular dimensions (*i.e.*, boundary condition at the bounding faces of the box, while in many cases these faces are within distances to the molecule comparable to the molecular size)

In this paper, we apply accurate methods to compute molecular shapes for use in BEM electrostatics calculations. We also determine the desired formulation of BEM which has the advantage of numerical stability. We show such numerical stability is scalable when more mesh elements are introduced under this formualation, and provides an important advantage for practical computation. Through an example, we demonstrate that this characteristics is not shared by the other widely used BEM formulations for molecular electrostatics. We further report developments of effective methods for numerical integration of singular and near singular integrands, which are essential for an accurate BEM calculation.

2 Poisson-Boltzmann Equaion and the Formulation of the Integral Equations

For a molecule Ω embedded in an ionic solution, its interior (int Ω) and the exterior ($\mathbb{R}^3 - \Omega$) are separted by the surface of the molecule, *i.e.*, the boundary $\partial\Omega$ of the molecule. For any three dimensional point $x \in \mathbb{R}^3$ in the interior of the molecule, the electrostatic potential $u^i(x)$ is governed by the Poisson equation:

$$\nabla^2 u^i(x) = -\frac{1}{\epsilon_i} \sum_{k=1}^{n} q_k \cdot \delta(x - x_k), \quad \text{for } x \in \text{int } \Omega \tag{1}$$

Here ϵ_i is the uniform dielectric constant inside a molecule, $\delta(x)$ is the three-dimensional delta function, and the interior of the molecule contains n point-charges q_k at position x_k. For any point in the exterior $x \in \mathbb{R}^3 - \Omega$, the

potential $u^e(x)$ is governed by the linearized Poisson-Boltzmann equation:

$$\nabla^2 u^e(x) = \kappa^2 \cdot u^e(x) \quad \text{for } x \in \mathbb{R}^3 - \Omega \tag{2}$$

Here $1/\kappa$ is the Debye length characterizing the screening effect due to the presence of the ions in the solvent. On the surface of the molecule $\partial\Omega$, we denote the normal derivative of the potential at x along the outward unit normal vector n_x at x as g, (*i.e.*, $g^i(x) = \partial u^i(x)/\partial n_x$ and $g^e(x) = \partial u^e(x)/\partial n_x$), then we have the interface condition:

$$u^i(x) = u^e(x) \quad \text{and} \quad g(x) \equiv g^i(x) = \frac{\epsilon_e}{\epsilon_i} \cdot g^e(x) \tag{3}$$

The above equations can be formulated into integral equations by standard BEM technique [17]. Two different formulations have been reported and details are described in [11, 12]. We call them *the non-derivative method* (nBEM) [10, 11, 18] and *the derivative method* (dBEM) [12]. In this section we briefly outline the main steps in the formulations and summarize the resulting equations and formulae.

nBEM Formulation. Applying Green's second identity

$$\int_V (\phi \nabla^2 \psi - \psi \nabla^2 \phi) \cdot dV = \int_S (\phi \nabla \psi - \psi \nabla \phi) \cdot dS \tag{4}$$

to the Poisson equation (1), we can obtain the integral equation for potential in the interior:

$$\omega(x) \cdot u^i(x) = \int_{\partial\Omega} K_1(x,y) \cdot g(y) \cdot d\sigma_y - \int_{\partial\Omega} K_2(x,y) \cdot u(y) \cdot d\sigma_y + \sum_{k=1}^{n} \frac{q_k}{4\pi\epsilon_i |x - x_k|}, \quad x \in \text{int } \Omega, \quad y \in \partial\Omega. \tag{5}$$

Here x is the location of an interior point, y is a surface point, and $\omega(x)$ is the solid angle at point x. For interior $x \in$ int Ω, $\omega(x) = 1$ (the full angle is counted as 1). The integration is over all points on the surface of the molecule. K_is are kernel integrands and are explained later.

Similarly applying Green's second identity to the linearized Poisson-Boltzmann equation (2), we obtain the integral equation for the exterior potential:

$$\omega(x) \cdot u^e(x) = -\frac{\epsilon_e}{\epsilon_i} \int_{\partial\Omega} K_3(x,y) \cdot g(y) \cdot d\sigma_y \quad + \quad \int_{\partial\Omega} K_4(x,y) \cdot u(y) \cdot d\sigma_y, \quad x \in \mathbb{R}^3 - \Omega, \quad y \in \partial\Omega. \tag{6}$$

Again here $\omega(x) = 1$.

For surface points $x \in \partial\Omega$, the solid angle for the above two equation is no longer the full angle. For example, if point x is in a smooth neighborhood, then $\omega(x) = 1/2$. Combining integral equations Eqn. 5 and Eqn. 6 with the interface condition Eqn. 3, we obtain the boundary integral equations for surface potential and its normal derivative:

$$\begin{aligned} \frac{1}{2} \cdot u(x) - \int_{\partial\Omega} K_1(x,y) \cdot g(y) \cdot d\sigma_y \quad + \quad & \int_{\partial\Omega} K_2(x,y) \cdot u(y) \cdot d\sigma_y \\ & = \sum_{k=1} \frac{q_k}{4\pi\epsilon_i |x - x_k|} \end{aligned} \tag{7}$$

$$\begin{aligned} \frac{1}{2} \cdot u(x) + \frac{\epsilon_e}{\epsilon_i} \int_{\partial\Omega} K_3(x,y) \cdot g(y) \cdot d\sigma_y \quad - \quad & \int_{\partial\Omega} K_4(x,y) \cdot u(y) \cdot d\sigma_y \\ & = 0 \end{aligned} \tag{8}$$

The integrand kernels in these integral equations are:

$$\begin{aligned} K_1(x,y) &= \frac{1}{4\pi|x-y|}, \qquad K_2(x,y) = \frac{1}{4\pi} \cdot \frac{\partial}{\partial n_y}\left(\frac{1}{|x-y|}\right), \\ K_3(x,y) &= \frac{e^{-\kappa|x-y|}}{4\pi|x-y|}, \qquad K_4(x,y) = \frac{1}{4\pi} \cdot \frac{\partial}{\partial n_y}\left(\frac{e^{-\kappa|x-y|}}{|x-y|}\right) \end{aligned} \tag{9}$$

where K_1 is the free-space Green's function (or the fundamental solution) to the Poisson's equation (1), K_3 the free-space Green's function (or the fundamental solution) to the linearized Poisson-Boltzman equation (2). K_3 and K_4 are their respective directional derivative along the outward normal vectors at y. The earlier integrals can thus

be thought as integrating over a single layer point charge distribution (K_1 and K_3 of charge density $g(y)$, without and with screening respectively) and a double layer point dipole distribution (K_2 and K_4 of dipole density $u(y)$, without and with screening respectively) on the surface of the molecule.

dBEM Formulation. The first equation for surface point $x \in \partial\Omega$ of this formulation can be obtained by combining the two boundarly integral equations Eqn. 7 and Eqn. 8 in the nBEM formulation, with consideration of the interface condition Eqn. 3. The second equation can be obtained by taking the normal derivative along the direction of the outward normal vector at x. For a surface point x in a smooth neighborhood, these two equations can be written as [12]:

$$\begin{aligned} \frac{1}{2}(1+\frac{\epsilon_e}{\epsilon_i})\cdot u(x) - \int_{\partial\Omega} K_1(x,y)\cdot u(y)\cdot d\sigma_y \quad &+ \quad \int_{\partial\Omega} K_2(x,y)\cdot g(y)\cdot d\sigma_y \\ &= \sum_{i=1} \frac{q_k}{4\pi\epsilon_i|x-x_k|} \end{aligned} \tag{10}$$

$$\begin{aligned} \frac{1}{2}(1+\frac{\epsilon_i}{\epsilon_e})\cdot g(x) - \int_{\partial\Omega} K_3(x,y)\cdot u(y)\cdot d\sigma_y \quad &+ \quad \int_{\partial\Omega} K_4(x,y)\cdot g(y)\cdot d\sigma_y \\ &= \sum_{i=1} \frac{q_k}{4\pi\epsilon_i}\cdot\frac{\partial}{\partial n_x}(\frac{1}{|x-y|}) \end{aligned} \tag{11}$$

where the kernels are now:

$$\begin{aligned} K_1(x,y) &= \frac{1}{4\pi}\cdot\frac{\epsilon_e}{\epsilon_i}\cdot\frac{\partial}{\partial n_y}(\frac{e^{-\kappa|x-y|}}{|x-y|}) - \frac{1}{4\pi}\cdot\frac{\partial}{\partial n_y}(\frac{1}{|x-y|}) \\ K_2(x,y) &= \frac{1}{4\pi}\cdot(\frac{1}{|x-y|} - \frac{e^{-\kappa|x-y|}}{|x-y|}) \\ K_3(x,y) &= \frac{1}{4\pi}\cdot\frac{\partial^2}{\partial n_x\partial n_y}\cdot\frac{e^{-\kappa|x-y|}}{|x-y|} - \frac{1}{4\pi}\cdot\frac{\partial^2}{\partial n_x\partial n_y}\frac{1}{|x-y|} \\ K_4(x,y) &= \frac{1}{4\pi}\cdot\frac{\partial}{\partial n_x}(\frac{1}{|x-y|}) - \frac{1}{4\pi}\cdot\frac{\epsilon_i}{\epsilon_e}\cdot\frac{\partial}{\partial n_x}(\frac{e^{-\kappa|x-y|}}{|x-y|}) \end{aligned} \tag{12}$$

Using these kernels, the interior and exterior potentials are:

$$u^i(x) = \int_{\partial\Omega} K_1(x,y)\cdot u(y)\cdot d\sigma_y + \int_{\partial\Omega} K_2(x,y)\cdot g(y)\cdot d\sigma_y + \sum_{k=1}^{n} \frac{q_k}{4\pi\epsilon_i|x-x_k|}, \quad x \in \text{int } \Omega, \quad y \in \partial\Omega. \tag{13}$$

$$u^e(x) = \frac{\epsilon_i}{\epsilon_e}\int_{\partial\Omega} K_1(x,y)\cdot u(y)\cdot d\sigma_y + \frac{\epsilon_i}{\epsilon_e}\int_{\partial\Omega} K_2(x,y)\cdot g(y)\cdot d\sigma_y + \sum_{k=1}^{n} \frac{q_k}{4\pi\epsilon_e|x-x_k|}, \quad x \in \mathbb{R}^3 - \Omega, \quad y \in \partial\Omega. \tag{14}$$

The regularity assumptions and the limiting process are argued in [12]

Numerical Method to Solve the Boundary Integral Equations. For numerical solution, we partition the surface of the molecule $\partial\Omega$ into a set of disjoint planar triangles Γ_i, so $\partial\Omega = \bigcup\Gamma_i$. We then approximate the unknowns $u(x)$ and $g(x)$ over a triangular element Γ_i by linear interpolation of the nodal values of u and g of this triangle, using linear basis (shape) functions and barycentric coordinates. The nodal values of n and g are then the unknowns we seek. The integrals over the whole surface become the summation of elemental integration, which are carried out by numerical quadrature. Seeking the solutions to the nodal values of u and g, we obtain $2M$ linear equations for a molecular mesh consisting of M vertices. Conjugated gradient based iterative solver (such as GMRES and BICGSTAB) is then employed to solve these equations. Details of various numerical set up of BEM can be found in [19].

3 Accurate Representation of the Molecular Shape

Under the continuum approach, different forms of the partial differential equations (PDE) have been applied to model the molecular electrostatics : Poisson equation [6, 8, 14], linearized Poisson-Boltzmann equation [20, 7, 12], and the non-linear Poisson-Boltzmann equation [21], each applicable under specific physical condition (such as ionic strength). Of equal importance to the nature of the elliptic differential equation is the boundary descriptor, *i.e.*, the accurate representation of the shape of the molecule. An exact molecular shape description is therefore crucial

for an accurate solution to the molecular electrostatics problem using the continuum model.

Molecular Surface Mesh Generation with Topological Precision. An important development for geometric modeling of macromolecules is the application of Delaunay complexes and the related alpha shape theory of alpha complexes. Alpha shape theory provides a topologically precise method for analytical computation of the molecular shape. In this method, the weighted Delaunay triangulation of the atom centers of the molecule is first computed [22]. The Delaunay complex obtained consists of tetrahedra, triangles, and edges connecting the atom centers, and vertices representing the atom centers. The alpha complex is then obtained by sorting the elements using a ball-growth model [23]. In addition to the vertices, the alpha complex retains a subset of the Delaunay tetrahedra, triangles, and edges, each corresponds to an occurrance of overlapping of the atom balls. The molecule and the alpha complex correspond to each other at topological, combinatorial and metric levels. Algorithms have been developed to compute analytical molecular shapes for molecules, including the calculation of analytical area/volume of solvent accessible (SA) model and molecular surface (MS) model [24], identification and measurements of inaccessible cavities in molecules [25], and quantitative computation of the molecular pockets and analytical construction of receptor binding sites [26].

In this work, we employ an alpha shape-based method to topologically triangulate the surface of the molecule [27]. The high quality mesh generated allows precise representation of the molecular shape, and maintains convenient data structure for easy access to geometric and topological information contained in the mesh. Detailes of the mesh generation can be found in [27].

Analytical Representation of Curve-Shaped Panel. For boudnary integral equation based BEM methods, the representation of the surface $\partial\Omega$ based on the mesh of planar triangles is important: curved panels are necessary to accurately represent both the shape of the surface integration domain and to evaluate the kernels. All interior points of the planar triangles therefore need to be mapped onto the atomic spheric surfaces.

Instead of using a quadratic or cubic function of the local coordinates in a planar triangle to approximate the

spherical surface [10, 12], here we describe an analytically exact method to map a planar point to the spherical patch. We exploit the fact that in the solvent accessible surface model, the surface of the molecule consists of pieces of surface patches with known analytical expressions. Our representation is piecewise C^∞ (differentiable to degree infinity in the interior of each patch), and therefore is faithful to the original geometric model of the molecule [28]. This method builds on the fact that geometric information is readily available from the data structure of the alpha shape generated mesh.

In the SA model, we first compute the vector $v \in \mathbb{R}^3$ from the center of the atom to the planar point $r \in \mathbb{R}^3$. Normalization of v will give the unit vector normal $n(r') = v/||v||$ to the surface point $r' \in \mathbb{R}^3$ to be mapped, which is on the spheric surface. We obtain the spheric point r' by:

$$r' = r + f \cdot n(r')$$

where $f = (R - ||v||)/||v||$, R being the radius of the atom. (See Figure 1). The Jacobian determinant of this transformation is $|J(r')| = 1/|k(r) \cdot n(r')|$, where k is the normal vector of the planar triangle. Geometrically, it is the ratio of the area of an infinitesimally small surface piece and the area of its projection on the planar triangle.

4 Numerical Quadrature

The formulae for the kernels of the boundary integral equations (Eqns. 9 and 12) reveal a difficulty in their integrations: these integrands become singular when y approaches x. The integral of these kernels all have interpretations: these kernels are either weakly singular, or have Cauchy principal values, or their finite parts in the Hadamard sense exist [29, 30]. For kernel $K_3(x, y)$ in Eqn. 12, each of the term is hypersingular, but it was shown that $K_3(x, y)$ as their difference is integrable and well behaved [12].

In BEM, these singularities hamper the numerical integration when x and y are on the same element (singular

case), or when they are very close (near singular case). Any quadrature rule of fixed weights (such as increasing the number of Gaussian quadrature points) is inadequate for the singular case, because these input-sensitive integrals are strongly influenced by the exact local shape of the element [31]. For the near singular case, although the integrals are regular, their numerical behavior is sensitive: Gauss formula loses its asymptotic accuracy because the estimated error is influenced by the derivative of the kernels, which produce negative powers of the stepsize, cancelling the asymtotic gain from using a higher order formula [32]. In the molecular electrostatic problem, previous applications of BEM often use Gaussian quadrature for singular integrands, and the problem of near singular integrands has been overlooked.

The accurate evaluation of the integration of the kernels, multiplied by the shape (basis) fuctions is one of the key issues in the BEM, since the resulting singular diagonal elements of the matrix strongly affect the whole solution. Although analytical formulae are possible for planar triangle elements, integration over curved geometric elements is only plausible with numerical quadrature rules. In numerical analysis, the quadratures for singular integrands and near-singular integrands have been a very active research field in the past decade [33, 34, 31, 32].

In our approach, we map all planar triangles Γ_i to a standard simplex or a master triangle $(0.0, 0.0; 0.0, 1.0; 1.0, 0.0)$ where the singular point is mapped to $(0.0, 0.0)$. Numerical integration is then carried out over this standard triangle for all patches of the surface. The mapping from a planar triangle to a spheric trianglular patch is as described earlier. Our method of numerical integration over the planar triangle is adapted from [34] , where error analysis of several classes of kernels are available for curve-shaped elements.

Singular Integrand. Briefly, for the radial variable, we use a polar transformation to weaken the singularity. The new variable $\xi_1 \in [0, 1]$ is chosen such that:

$$d\xi_1 = \frac{dr}{r(\phi)} \tag{15}$$

In the master triangle, we further express $r(\phi)$ as:

$$r(\phi) = \frac{h}{\sin(\alpha + \phi)},$$

where h is the height of the triangle from the side opposite to the singular node, and $\alpha = \pi/4$ (See Figure 2).

For the azimuth angle variable, there are still large variations in $r(\phi)$ when angles are near either of the two edges to which the singular point is the common endpoint. The situation is worse when the angle apexed by the singular point in the underlying Γ_i is obtuse. Numerically this introduces undesirable instability. A further transformation for the azimuth angle variable is therefore necessary, such that the quadrature sampling points can be distributed unevenly. To take into account the large variation with the angle variable, the sampling points in the vicinity of the two edges should be finer than in between the edges. The measure adapted to this situation is $d\phi/\sin\phi$ rather than $d\phi$. Modified from [34], we use the new variable $\xi_2 \in [-1, +1]$ for the azimuth angle variable, such that:

$$d\xi_2 = \frac{2 \cdot d\phi}{\Delta_k \cdot \sin(\phi + \alpha)}, \tag{16}$$

where

$$\Delta_k := \int_{\alpha}^{\alpha + \frac{\pi}{2}} \frac{d\theta}{\sin\theta + \cos\theta} = 2\ln(1 + \sqrt{2}), \quad \alpha = \frac{\pi}{4}.$$

From formulae (15) and (16), and the ranges of ξ_1 and ξ_2, we have the following transformation of the polar coordinates to the new variables ξ_1 and ξ_2 to be used in the Gaussian quadrature:

$$\begin{aligned} \xi_1 &= r\frac{\sin(\alpha + \phi)}{h}, \\ \xi_2 &= \frac{1}{\Delta_k}(2\ln\tan\frac{\phi + \alpha}{2} - \ln\tan\frac{\alpha}{2} - \ln\tan\frac{\pi/2 + \alpha}{2}) \end{aligned}$$

The tranformations from the new variables ξ_1 and ξ_2 back to $r(\phi)$ and ϕ are:

$$\begin{aligned} \phi &= 2 \cdot \arctan(e^{\xi_2 \cdot \ln(1+\sqrt{2})}) - \frac{\pi}{4} \\ r &= \frac{\xi_1 \cdot h}{\sin(\phi + \alpha)} = \frac{\xi_1}{\sqrt{2}\sin(\phi + \frac{\pi}{4})}, \end{aligned}$$

Applying Gaussian quadrature to both ξ_1 and ξ_2 variables, we have the following quadrature rules:

$$\int_{\Gamma_i} K(x,y) \cdot u(y) d\sigma_y \approx \frac{\ln(1+\sqrt{2})}{\sqrt{2}} \sum_{n=1}^{N} \sum_{l=1}^{L} w_{2n} w_{1l} \cdot r_{ln} K(x, r_{ln}, \phi_n) \cdot u_i(r_{ln}, \phi_{ln}) \cdot J_i(r_{ln}, \phi_{ln}) \tag{17}$$

where $r_{ln} := r_l(\phi_n) = r_l(\xi_{2,n})$, $\phi_n = \phi(\xi_{2,n})$, and N, L are the numbers of quadrature points of transformed variables ξ_1 and ξ_2.

Nearly Singular Integrand. One remedy for the near singular integrand is to expand the unknown functions with Taylor series into a known sharp changing function and an unknown smooth varying function. This is roughly the "peak splitting" method used by Juffer *et al.* in potential evaluation (although it is not used in influence matrix assembly) [12]. However, the Coulomb term dominates the Taylor expansion, since the majority of the atoms carry partial charges and cannot be treated as neutral. Hence, the evaluation of the known function at one point in space will take $O(n)$ steps to sum over all charges. The surface integration (itself an $O(n^2)$ process, if all values of the function are known) will render the overall performance of the algorithm close to $O(n^3)$ time complexity for computing potential at one point in space. In the interest of practical time complexity, we do not follow this approach. Instead, we follow [34]: when x is near the panel Γ_i, the standard simplex or the master triangle is subdivided into P^2 geometrically similar sub-simplexes. We then apply Gaussian quadrature to each of these sub-simplexes. In our implementation, we measure the distance $|x - y|$, and apply this scheme if the distance is smaller than the longest edge of the planar triangle. Details of error analysis of this rule can be found in [34].

5 Results

Conditioning of the Boundary Integral Equations with dBEM. Conceptually, boundary integral equations lead to a dense, nonsymmetric influence (coefficient) matrix for a system of linear equations to be solved numerically. In principle, Poisson equation leads to a second kind Fredholm type integral equation, characterized by a well-conditioned compact integral operator. This significantly facilitates the solution of the resulting set of linear equations, and allows effective use of a class of iterative solvers [35]. In our cases, the non-integral terms on the left hand side of the boundary integral equations of dBEM distinguish these equations as the second kind Fredholm type integral equations. We use increasing number of triangular mesh elements to model a unit sphere, and we calculate the condition number of the resulting influence matrices. Condition number of a matrix A ($cond(A) = ||A|| \cdot ||A^{-1}||$) measures how sensitive the solution of a linear system is to perturbation. It is a good indicator of the convergence behavior of many iterative solvers for linear equations, where floating point calculations are always accompnied with errors from finite arithmetic [35]. In all cases, the matrices of the second kind integral equations resulting from dBEM have significantly lowered condition number, as compared to the matrices where the non-integral terms are left out.

However, the matrices from nBEM and dBEM have very different characteristics. This can be seen from the different forms of Eqns. (7,8) and (10, 11), and from the visualization of the coefficient matrices. Figure 3 plots the coefficient matrices from both the nBEM and the dBEM formulations for a 16-vertex triangulation of a unit sphere. This suggests that nBEM and dBEM formulations may have different convergence behavior for many numerical solvers. As an experiment, we use 16, 32, 64, 128 and 256-vertex triangulation of the unit sphere to test how the condition number scales with the size of the matrices from the two formulations. Sphere has the simplest geometry, and there is no effect due to complicated geometric considerations. As a result, such a test should reflect the characteristics of matrices due to the choice of formulation and exclude other factors. Figure 4 shows that the condition numbers of the dBEM matrices (empty circle) are constantly small as the size of the matrix increases from 32 (for 16 vertex) to 512 (for 256 vertices). For nBEM (filled circle), the condition number deteriorates rapidly as the size of the matrix increases.

This has implications because rapid convergence of the linear equation solvers is an important issue in BEM, when meshes of large sizes are necessary to model molecules with compelx geometry. There is no difference in terms of storage required for both dBEM and nBEM matrix. Therefore, we suggest that it is advantageous to use the dBEM formulation for a better conditioned influence matrix.

Numerical Tests. As a test case, we computed the electrostatic reaction potential of a unit point charge placed at various eccentric locations inside the unit ball, where the analytical solutions are known. We set the protein interior with an ϵ_i of 1, and the continuum solvent of ϵ_e of 20 outside the ball. The inverse Debye-Huckel length κ is set to 3/Å. For these calculations, dBEM formulation is used. An iterative solver (BICGSTAB) is used for solving the resulting linear equations [36] with the convergence residue set to 1.0×10^{-8}. The results are listed in Table 1. The improvement of accuracy of the solution can be seen as the number of vertices used for triangular elements increases. The iteration numbers for BICGSTAB solver indeed are small and shows little increase as the number of vertices used to triangulate the sphere increased from 64 to 512. This agrees with our earlier assessment based on condition number calculations.

For comparison, we have also implemented the nBEM formulation for computing the reaction potential for unit ball. The number of iterations using BICGSTAB solver is listed in Table 2 for the nBEM method. Compared to dBEM method, the iteration numbers are larger for all levels of discretization. The nBEM iteration numbers also increase as the discretization becomes finer, and this is in contrast to dBEM method. These results demonstrate that dBEM can provide accurate solutions with a better-conditioned influence matrix. This is reflected by the fact that in the examples studied, the number of iterations required to achieve comparable accuracy is invariant to the mesh size.

Figure 5 shows the electrostatic potentials computed through dBEM at the y-z plane of x = 8.0 Åwhen an ammonium molecule (NH_3) is placed in an infinit solution, centering at the origin and with its N atom along the z-axis. The $\epsilon_i$and $\epsilon_e$are set to be 2.0 and 80.0 respectively. The temperature and the monovalent ionic strength are set to be 298 K and 150.0 mM. The molecule is triangulated by 492 elements. Figure 5(A) shows the potential

surface on the x-y plane, and (B) shows the projected contour plot. In NH_3, N atom has a negative partial charge and H atoms have positive partial charges. The y-z plane has an asymmetric pattern with the high z part negative and the low z part positive, consistent with the locations of negative N (z = 0.110) and postive H (z = -0.256) atoms. Figure 6 shows the potentials computed through dBEM at the x-y plane of z = 8.0 Å. In contrast, the pattern of the potentials are rather symmetric, agreeing the geometric symmetry of the molecule.

As a preliminary study of protein electrostatics , we compute the potentials near a crambin molecule on the x-z planes at fixed y = -10 Å beneath the molecule, and at y = 26 Å above the molecule. The crambin molecule placed in an infinite medium has all its atom centers located in the following coordiate ranges: $x \in [-3.51, 25.02], y \in [-0.967, 20.867], z \in [-7.383, 19.238]$ and is centered at (-1.407, -0.684, -3.911). Figure 7 visualizes the potentials and its contour for y = -10 Å plane, and Figure 8 viusalizes the potentials and its contour for y = 26 Å plane.

6 Discussion

BEM differs from FDM in that it allows the possiblity to maintain an accurate (implicit) global solution using exact shape representation of the molecule [15]. FDM relies on technique such as "focusing" to obtain accurate local solutions. The alpha shape based mesh generation allows very accurate description of the molecular shape for BEM calculation, where no erroneous triangles will be added to connect across clefts and surface invaginations. It also faithfully represents the topolgoical features of the molecule, such as all the inaccessible cavities inside the molecule, channels, etc.

The kernel singularities of the boundary integral equations are one of the distinctive charcateristics of the BEM method. The strength of the singularities strongly affects the condition number of the resulting coefficient matrix. Several methods have been described to approach the singularity problem. Yoon and Lenhoff [11] analytically integrated the kernels for the nBEM formulation. However, in general the analytical evaluation of singular kernels [30] can be prohibitive in terms of algebraic derivation: it is only plausible when planar boundary elements are

used (as in [11]), with the price of introduced error due to inappropriate shape representation.

When ion effects are ignored from the continuum model, Poisson equation is used to model the molecular electrostatics. For such cases, Purisima and Nilar used "row-sum elemination" method to obtain the singular integration term [37]. It is based on the fact that the solid angle subtended by a point on a smooth surface is 1/2 of the full solid angle (4π). This method cannot be extended for two reasons. First, for the linearized Poisson-Boltzmann equation where ion effects are included, no simple geometric formulae are known for kernels of the type in Eqns. 9 (K_3, K_4), and Eqns. 12 (K_1 to K_4) to be used for subtraction. Second, when potential inside each element is not treated as a constant, the solid angle formula cannot be applied, since the basis fuctions which are used to interpolate between nodal values cannot be separated out from the integrand. On the other hand, Guermond showed that non-constant interpolation of the potential within each element is necessary to achieve accurate solution, if the use of a modest number of mesh elements is desired [38].

The conditioning of the boundary element equations is an important and complicated issue. We have shown that the dBEM formulation has an advantage over nBEM formulation: it results in a better conditioned influence matrix, and this formulation alone maintains a matrix with a good condition number even when the problem scales up. The overall conditioning of the boundary element equations is necessarily a complicated issue, and depends on other factors, including the shape of the molecule, the discretization of the surface of the molecule, the choice of the shape (basis) functions for approximating potentials whithin each element, and the approximation of the surface. Together with the numerical integration scheme, these factors also affect the accuracy of the solution. Further investigations in these directions will provide valuable information.

There have been recent efforts using multipole expansion to rapidly compute either explicitly or implicitly the influence matrix for Poisson equation [14, 15]. However, currently such efforts have been confined to Poisson equation where modeling of the ion effects are not possible. The reason behind is that the multipole expansions of the Green's function (and its derivative) for linearized Poisson-Boltzmann equation (*e.g.,* K_3 and K_4 in Eqn. 9) have not been worked out. Nevertheless, future fast methods should also significantly benefit from a stable and

better-conditioned influence matrix obtainable from the dBEM formulation. This is particularly relevant if the state-of-art nonstationary conjugate gratdient based iterative solvers (*e.g.*,GMRES, BICGSTAB,*etc.*) are used rather than the stationary iterative solvers such as Gauss-Seidel Method or the Successive Overrelaxation Method (SOR).

In conclusion, since geometric shape description is important for the continuum approach to the molecular electrostatics problem, we apply an alpha shape based method to generate topologically precise and analytically accurate surface mesh for molecules. We also describe the shape of the molecule analytically within each element, using a piecewise analytical function to represent the shape of the surface used in BEM. Contrary to widely used nBEM formulation [10, 11, 18], we show that the dBEM boundary integral equations provides superior numerical stability and should be used when speedy convergency in solution of the linear equations is desired. Singular and near-singular kernel integrands can introduce large errors to the solution in BEM, and needs to be treated carefully. We have described an effective method of variable transformation to be used for numerical quadratures.

7 Acknowledgement

We are indebted to Prof. Herbert Edelsbrunner for the topological surface triangulation which enables our approach. We thank Prof. Faisal Saied for many insightful discussions and suggestions, N. Akkiraju for generating the molecular meshes, J. Bordner for help in implementing the iterative solver. JL thanks Lu Wang for helpful discussions. JL is supported by an NSF CISE post-doctoral fellowship, grant ASC 94-04900. SS acknowledge support by NSF, grant ASC 89-02829, and MCB 92-19619. We thank NSF Meta Center Allocation for providing the computational resources.

References

[1] M.K. Gilson, A. Rashin, R. Fine, and B. Honig. On the calculation of electrostatic interactions in proteins. *J. Mol. Biol.*, 183:503–516, 1985.

[2] Stuart A. Allison, J. Andrew McCammon, and Scott H. Northrup. Dynamics of macromolecular interactions. In Adi Eisenberg and Fred E. Bailey, editors, *Coulombic Interactions in Macromolecular Systems*, pages 216–231. American Chemical Society, 1986.

[3] A.A. Rashin, M. Iofin, and B. Honig. Internal cavities and buried wtares in globular proteins. *Biochemistry*, 25:3619–3625, 1986.

[4] K.A. Sharp and B. Honig. Electrostatic interactions in macromolecules: theory and application. *Annu. Rev. Biophys. Biophys. Chem.*, 19:301–332, 1990.

[5] J.D. Madura, M.E. Davis, M.K. Gilson, R.C. Wade, B.A. Luty, and J.A. McCammon. Biological applications of electrostatic calculations and Brownian dynamics simulations. In K.B. Lipkowitz and D.B. Boyd, editors, *Reviews in Computational Chemistry*, volume V, chapter 4, pages 229–267. VCH Publishers, New York, 1994.

[6] J. Warwicker and H.C. Watson. Calculation of the electric potential in the active site cleft due to α-helix dipoles. *J. Mol. Biol.*, 157:671–679, 1982.

[7] M.E. Davis and J.A. McCammon. Electrostatics in biomolecular structure and dynamics. *Chem. Rev.*, 90:509–521, 1990.

[8] R.J. Zauhar and R.S. Morgan. A new method for computing the macromolecular electric potential. *J. Mol. Biol.*, 186:815–820, 1985.

[9] R.J. Zauhar and R.S. Morgan. The rigorous computation of the molecular electric potential. *J. Comp. Chem.*, 9:171–187, 1988.

[10] R.J. Zauhar and R.S. Morgan. Computing the electric potential of biomolecules: application of a new method of molecular surface triangulation. *J. Comput. Chem.*, 11:603–622, 1990.

[11] B.J. Yoon and A.M. Lenhoff. A boundary element method for molecular electrostatics with electrolyte effects. *J. Comput. Chem.*, 11:1080–1086, 1990.

[12] A.H. Juffer, E.F.F. Botta, B.A.M. van Keulen, A. van der Ploeg, and H.J.C. Berendsen. The electric potential of a macromolecule in a solvent: a fundamental approach. *J. Comp. Phys.*, 97:144–171, 1991.

[13] H-X Zhou. Boundary element solution of macromolecular electrostatics: Interaction energy between two proteins. *Biophys. J.*, 65:955–963, 1993.

[14] A. Bharadwaj, A. Windemuth, S.Sridharan, B. Honig, and A. Nicholls. The fast multipole boundary element method for molecular electrostatics: An optimal approach for large systems. *J. Comput. Chem.*, 16:898–913, 1995.

[15] R.J. Zauhar and A. Varnek. A fast and space-efficient boundary element method for computing electrostatic and hydration effects in large molecules. *J. Comput. Chem.*, 17:864–877, 1996.

[16] R.E. Bruccoleri. Grid positioning independence and the reduction of self-energy in the solution of the poisson-boltzmann equation. *J. Comp. Chem.*, 14:1417–1422, 1993.

[17] C.A. Brebbia, J.C.F. Telles, and L.C. Wrobel. *Boudnary element techniques: theory and applications in engineering*. Springer-Verlag, Berlin, 1984.

[18] H.-X. Zhou. Brownian dynamics study of the influence of electrostatic interaction and diffusion on protein-protein association kinetics. *Bioph. J.*, 64:1711–1726, 1993.

[19] P.K. Banerjee and S. Walker. *Boundary element methods in Engineering Science*. McGraw-Hill, New York, 1981.

[20] I. Klapper, R. Hangstrom, R. Fine, K. Sharp, and B. Honig. Focusing of electri fields in the active site of Cu-Zn superoxid dismutase: effects of ionic strength and amino-acid modifications. *Proteins Struct. Funct. Genet.*, 1:47–59, 1986.

[21] M. Holst and F. Saied. Numerical solution of the nonlinear Poisson Boltzmann equation: Developing more robust and efficient methods. *J. Comput. Chem.*, 16:337–364, 1995.

[22] H. Edelsbrunner and N.R. Shah. Incremental topological flipping works for regular triangulations. In *Proc. 8th Ann. Sympos. Comput. Geom.*, pages 43–52, New York, 1992. ACM Press.

[23] H. Edelsbrunner and E.P. Mücke. Three-dimensional alpha shapes. *ACM Trans. Graphics*, 13:43–72, 1994.

[24] J. Liang, H. Edelsbrunner, P. Fu, P.V. Sudhakar, and S. Subramaniam. Analytical shape computing of macromolecules I. molecular area and volume through alpha shape. *Proteins*, submitted, 1996.

[25] J. Liang, H. Edelsbrunner, P. Fu, P.V. Sudhakar, and S. Subramaniam. Analytical shape computing of macromolecules II. Inaccessible cavities in proteins. *Proteins*, submitted, 1996.

[26] H. Edelsbrunner, M. Facello, and J. Liang. On the definition and the construction of pockets in macromolecules. In *Proc. 1st Pacific Sympos. Biocomputing*, Singapore, 1996. World Scientific.

[27] N. Akkiraju and H. Edelsbrunner. Triangulating the surface of a molecule. *Discrete Appl. Math.*, In press, 1996.

[28] B. Lee and F. M. Richards. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.*, 55:379–400, 1971.

[29] H.R. Kutt. The numerical evaluation of principal value integrals by finite-part integration. *Numer. Math.*, 24:205–210, 1975.

[30] A.C. Kaya and F. Erdogan. On the solution of integral equations with strongly singular kernels. *Q. Appl. Math.*, XLV:105–122, 1987.

[31] C. Schwab and W.L. Wendland. On numerical cubatures of singular surface integrals in boundary element methods. *Numer. Math.*, 62:343–369, 1992.

[32] W. Hackbusch and S.A. Sauter. On numerical cubatures of nearly singular surface integral arising in BEM collocation. *Computing*, 52:139–159, 1994.

[33] C.G.L. Johnson and L.R. Scott. An analysis of quadrature errors in second-kind boundary integral methods. *SIAM J. Numer. Anal.*, 26:1356–1382, 1989.

[34] J-L Guermond. Numerical quadratures for layer potentials over curved domains in $\mathbb{R}^3$. *SIAM J. Numer. Anal.*, 29:1347–1369, 1992.

[35] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition*. SIAM, Philadelphia, PA, 1994.

[36] H.A. van der Vorst. Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of non-symmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 13:631–644, 1992.

[37] E.O. Pursima and S.H. Nilar. A simple yet accurate boundary element method for continuum dielectric calculation. *J. Comp. Chem.*, 16:681–689, 1995.

[38] J-L Guermond. High order numerical quadratures for layer potentials over curved domains in $\mathbb{R}^3$. *Comput. Methods Appl. Mech Engrg*, 116, 1994.

Table Legend

Table 1. Computed reaction potential due to a unit charge placed inside a unit ball at various eccentric locations. d is the distance of the charge from the center of the ball. The errors of the numerical results in % compared to analytical solutions are listed along with the iteration number of the BICGSTAB solver [36] with the convergence residue set to 1.0×10^{-8}. The inverse Debye-Huckel length κ is set to 3/Å. A 7×7 Gauss quadrature rule is used for the nonsingular integrand, a $10(\phi) \times 5(r)$ quadrature rule as described in the paper is used for the singular integrand. For the intermediate integrand, level 3 subdivision and 7×7 quadrature rule is used. Linear functions are used as the basis functions.

Table 2. Iteration numbers of BICGSTAB solver using nBEM formulation for computing reaction potential due to a unit charge placed inside a unit ball. All parameters are the same as in Table 1.

# vertices.		16		32		64		128		256		512	
d	Exact	Iter.	Error	Iter.	Error	Iter.	Error	Iter.	Error	Iter.	Error	Iter.	Error
0.0	-3.77184	5	4.29	7	2.39	7	1.28	8	0.67	8	0.35	6	0.18
0.1	-3.80928	9	4.43	10	2.45	10	1.31	10	0.68	11	0.36	10	0.18
0.2	-3.92544	9	4.87	10	2.62	11	1.39	12	0.72	12	0.37	10	0.19
0.3	-4.13616	10	5.78	10	2.98	12	1.55	13	0.80	13	0.41	11	0.21
0.4	-4.47250	11	7.62	12	3.61	13	1.84	14	0.93	16	0.48	13	0.25
0.5	-4.99440	14	11.38	12	4.79	16	2.35	17	1.16	17	0.58	16	0.30
0.6	-5.82816	13	18.74	12	7.21	15	3.38	15	1.66	16	0.80	16	0.42
0.7	-7.26768	13	31.51	13	12.81	15	5.79	17	2.92	15	1.33	17	0.73
0.75	-8.43504	14	40.21	14	18.06	15	8.26	17	4.41	16	1.92	18	1.06
0.80	-10.19376	13	50.34	15	26.03	16	12.61	16	7.45	19	3.11	21	1.66
0.85	-13.12656	14	61.65	16	37.64	19	20.55	18	13.82	21	6.00	22	2.88
0.90	-18.97824	13	73.84	16	53.56	20	35.07	20	27.20	22	14.00	22	5.69

Table 1:

d	It. #					
# vert.	16	32	64	128	256	512
0.0	5	15	18	22	24	25
0.1	10	17	19	22	32	33
0.2	10	21	19	27	31	37
0.3	11	21	23	26	29	36
0.4	12	20	21	31	41	37
0.5	14	24	27	36	37	50
0.6	14	23	30	46	52	43
0.7	15	25	40	57	62	45
0.75	17	36	36	50	58	58
0.80	16	34	41	47	57	56
0.85	19	34	42	47	62	62
0.90	21	40	39	57	75	65

Table 2:

Figure Legends

Figure 1. Mapping points on planar triangle to the surface of the atom.

Figure 2. h is the height of the triangle of the side opposite to the singular node. $\alpha = \pi/4$ is the side angle. Sampling points along the radial line $r(\phi)$ are of angle ϕ.

Figure 3. The matrices obtained from 16 vertices triangulation of a unit sphere using both nBEM and dBEM. Matrices are fromed using $\kappa = 3.0$, $\epsilon_i = 1$, and $\epsilon_e = 20$.

Figure 4. The condition numbers of the matrix formed by dBEM (empty circle) are constantly small as the matrix size increases, whereas condition number for nBEM matrices deteriorates rapidly. Matrices are fromed using $\kappa = 3.0$, $\epsilon_i = 1$, and $\epsilon_e = 20$.

Figure 5. The electrostatic potentials computed through dBEM on a y-z plane of x = 8.0 Å, when an ammonium molecule (ϵ_i= 2.0) centering at the origin is placed in an infinite monovalent ionic solution (150.0 mM, 298 K, ϵ_e= 80.0). The ammonium molecule is triangulated with 492 triangles. (A): the potential surface 8.0 Å below the center of the molecule along positive x direction (the inset shows the orientation of the amonium molecule); (B): the potential countour plot.

Figure 6. The electrostatic potentials computed through dBEM on a x-y plane of z = 8.0 Å of the above ammonium system. (A): the potential surface 8.0 Å below the center of the molecule along posistive z direction (the inset shows the orientation of the amonium molecule); (B): the potential countour plot.

Figure 7. The electrostatic potentials computed through dBEM on a x-z plane of y = -10.0 Å near a crambin molecule placed in an infinite medium. (A): the potential surface at y = -10.0 Å (the inset shows the orientation of the crambin molecule); (B): the potential countour plot.

Figure 8. The electrostatic potentials computed through dBEM on a x-z plane of y = 26.0 Å near a crambin molecule

placed in an infinite medium. (A): the potential surface at y = 26.0 Å (the inset shows the orientation of the crambin molecule); (B): the potential countour plot.

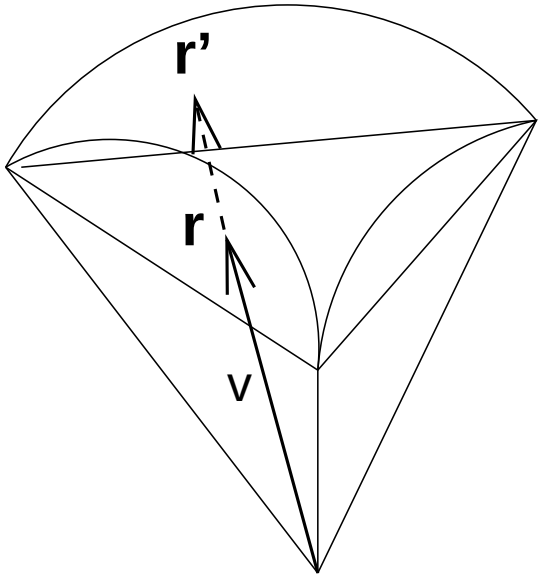


Figure 1:

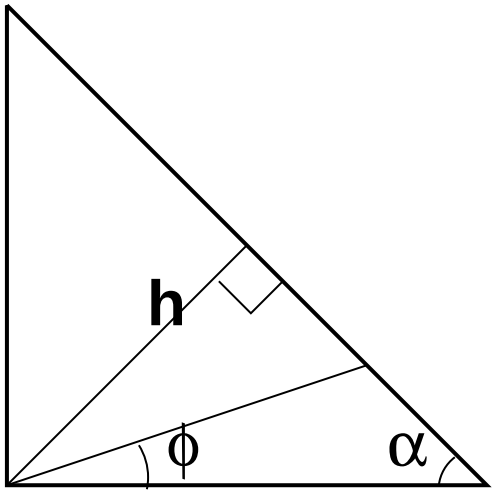


Figure 2:

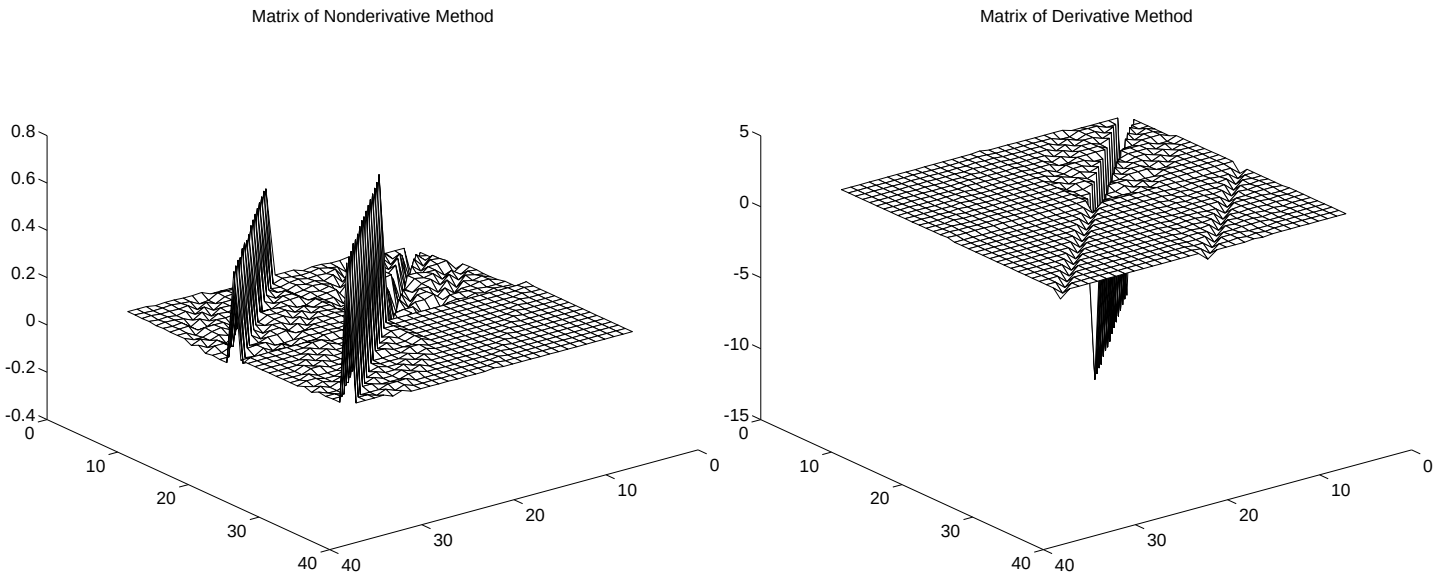

Matrix of Nonderivative Method
0.8
0.6
0.4
0.2
0
-0.2
-0.4
0
10
20
30
40
40
30
20
10
0
Matrix of Derivative Method
5
0
-5
-10
-15
0
10
20
30
40
40
30
20
10
0

Figure 3:

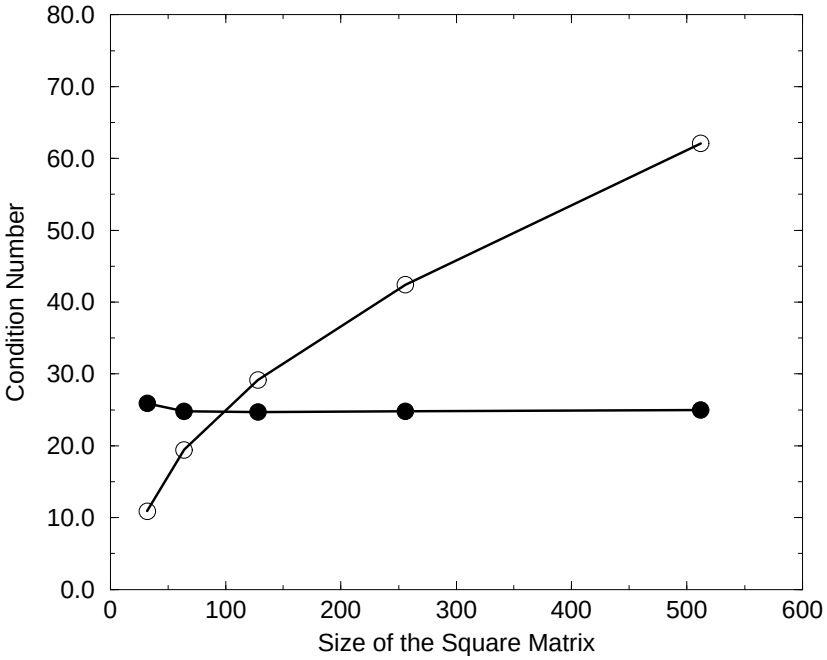
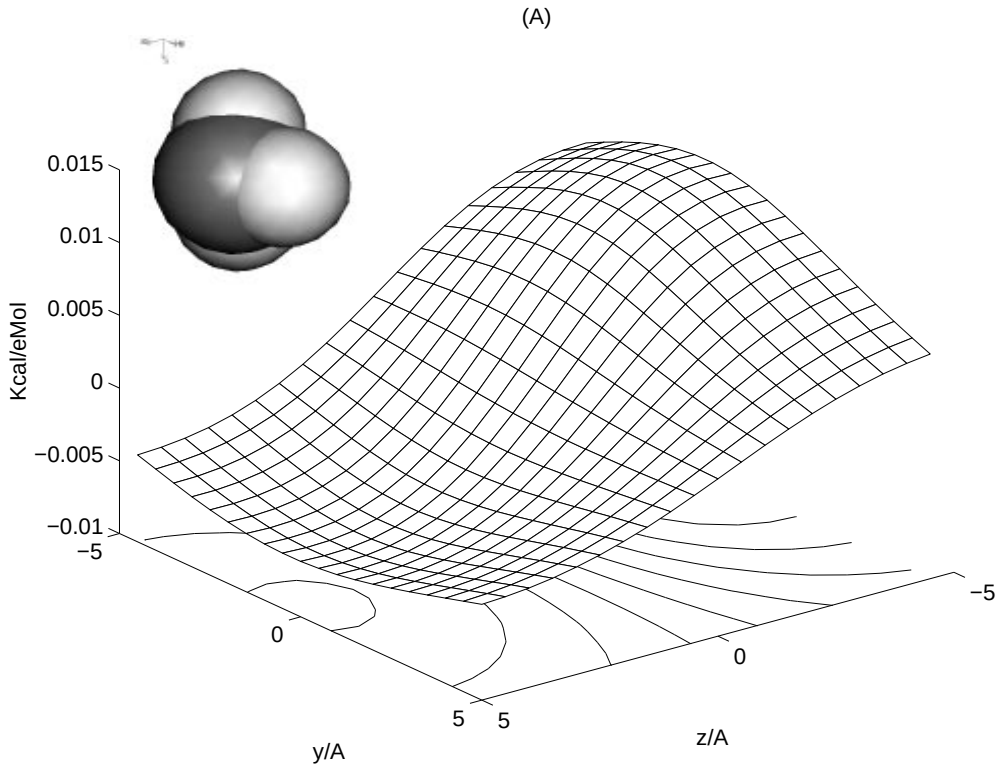

80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
Condition Number
0
100
200
300
400
500
600
Size of the Square Matrix

Figure 4:


(A)
0.015
0.01
0.005
0
−0.005
−0.01
Kcal/eMol
−5
0
5
y/A
5
0
−5
z/A

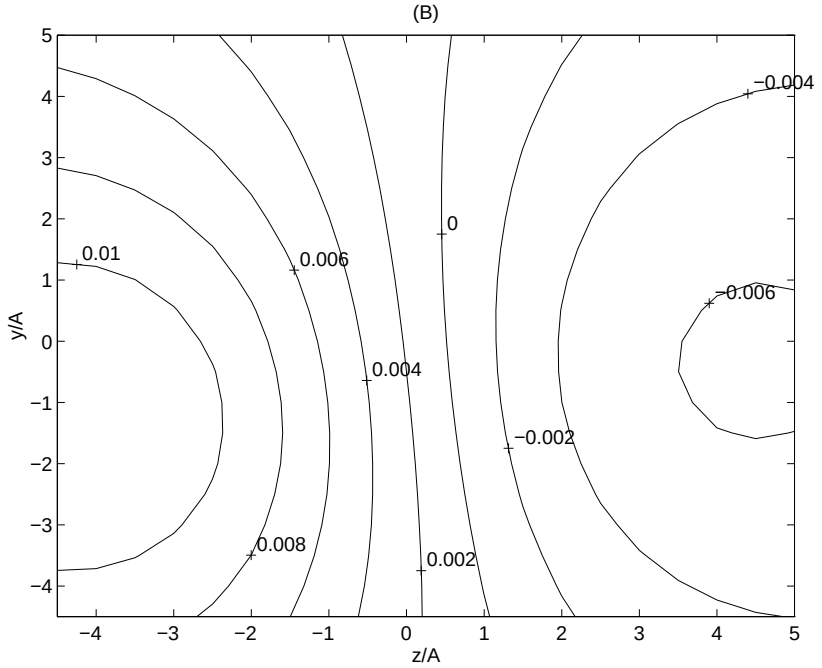
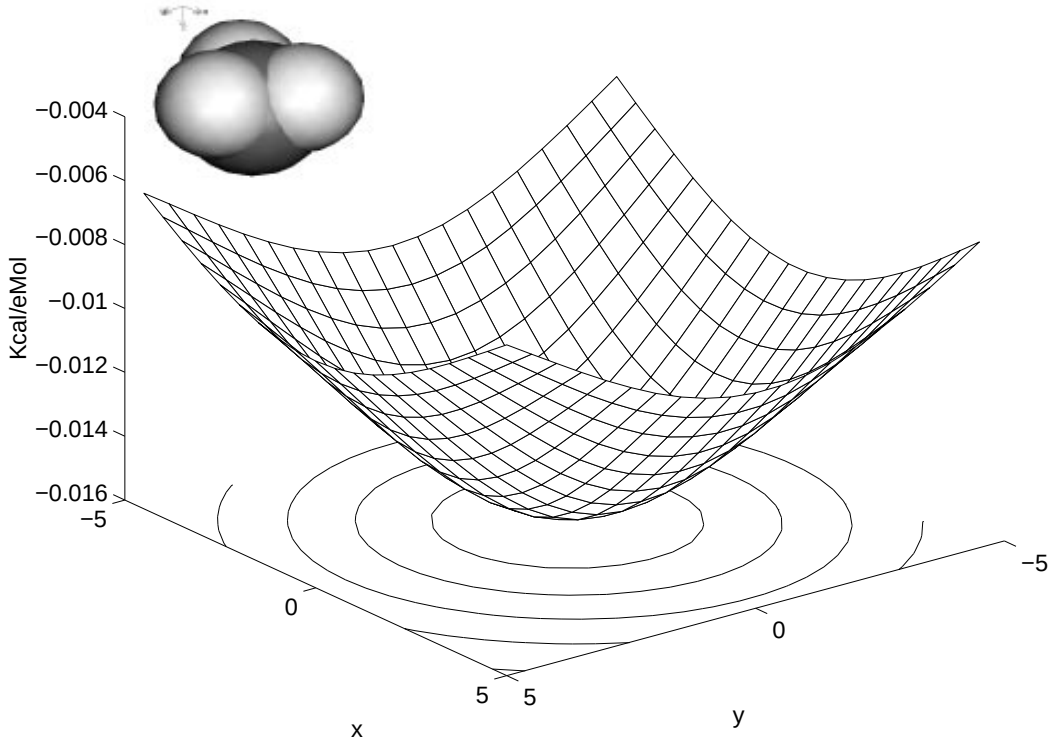

(B)
−0.004
0.01
0.006
0
0.006
0.004
−0.002
0.008
0.002
y/A
z/A

Figure 5:


−0.004
−0.006
−0.008
−0.01
−0.012
−0.014
−0.016
Kcal/eMol
−5
0
5
x
5
0
−5
y

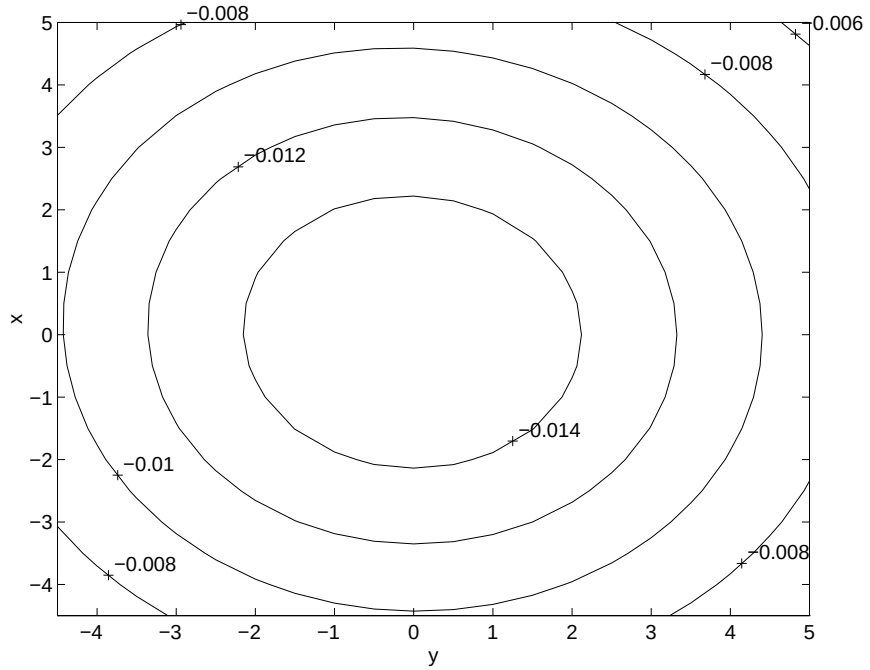
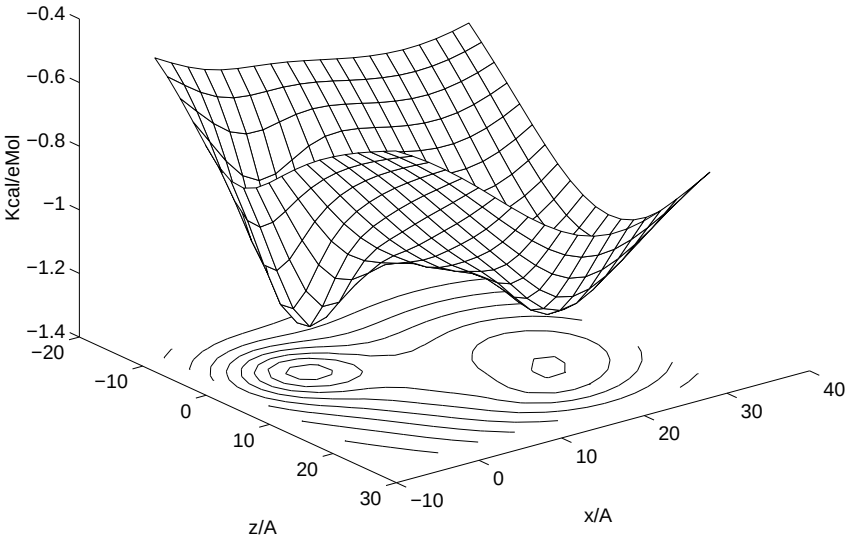

−0.008
−0.006
−0.008
−0.012
−0.014
−0.01
−0.008
−0.008
x
y

Figure 6:

(A)


−0.4
−0.6
−0.8
−1
−1.2
−1.4
Kcal/eMol
−20
−10
0
10
20
30
z/A
−10
0
10
20
30
40
x/A

(B)

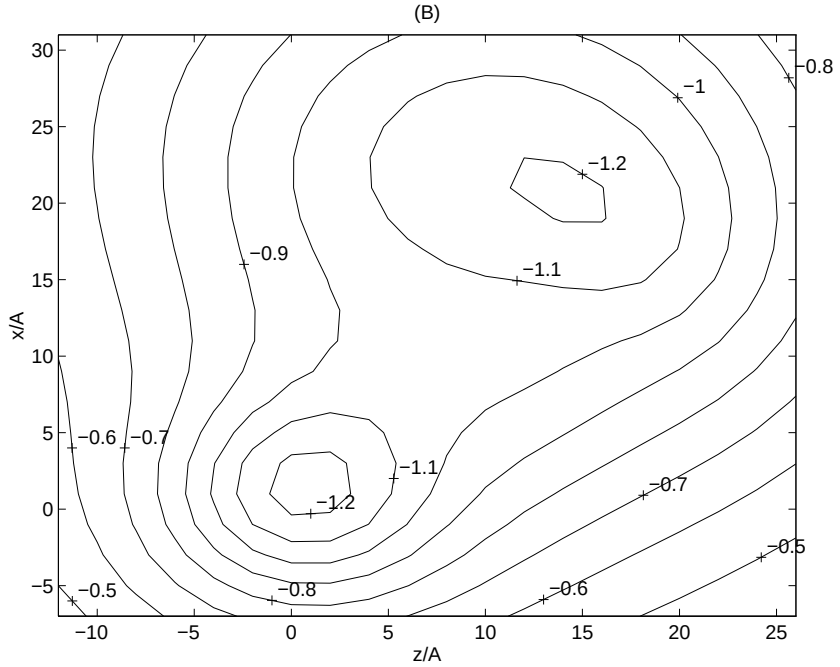
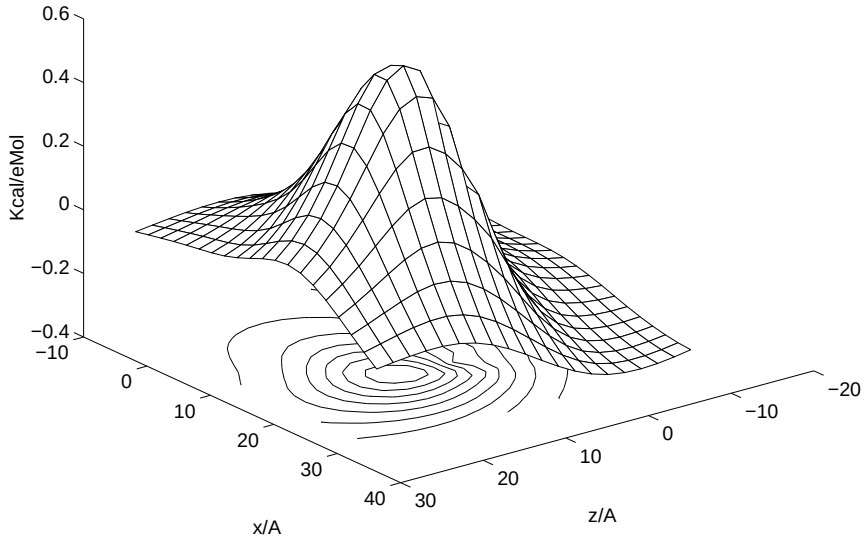

30
25
20
15
10
5
0
−5
x/A
−10
−5
0
5
10
15
20
25
z/A
−0.8
−1
−1.2
−0.9
−1.1
−0.6
−0.7
−1.1
−1.2
−0.7
−0.5
−0.5
−0.8
−0.6

Figure 7:

(A)


0.6
0.4
0.2
0
−0.2
−0.4
Kcal/eMol
−10
0
10
20
30
40
x/A
30
20
10
0
−10
−20
z/A

(B)

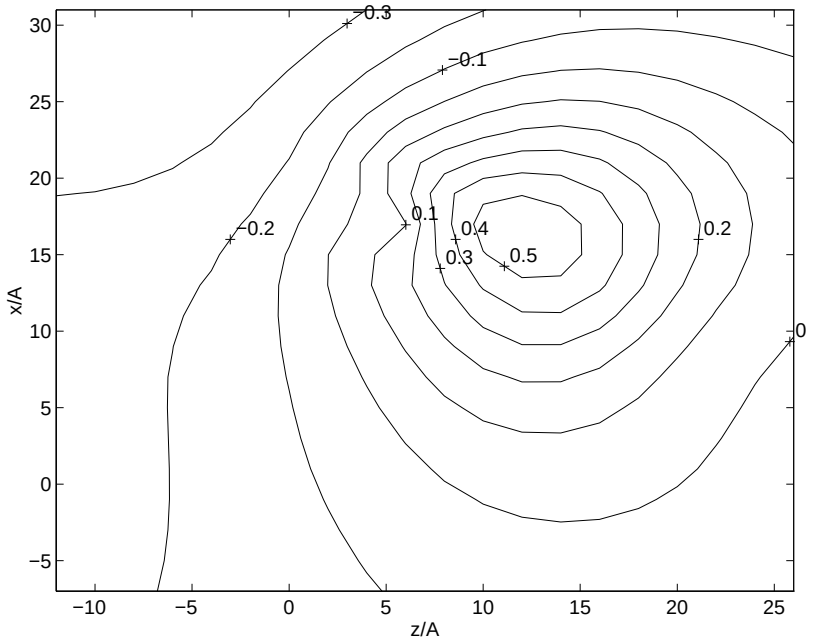

−0.3
−0.1
−0.2
0.1
0.4
0.3
0.5
0.2
0
30
25
20
15
10
5
0
−5
x/A
−10
−5
0
5
10
15
20
25
z/A

Figure 8: