Electrostatics of charged dielectric spheres with application to biological systems

T. P. Doerr* and Yi-Kuo Yu†
National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, 8600 Rockville Pike MSC 6075, Bethesda, Maryland 20894, USA
(Received 12 January 2006; published 2 June 2006)

Because electrostatic forces are crucial in biological systems, molecular dynamics simulations of biological systems require a method of computing electrostatic forces that is accurate and rapid. We propose a surface charge method, apply it to a system of arbitrary number of charged dielectric spheres, and obtain an exact solution for an arbitrary configuration of the spheres. The precision depends only on the number of terms kept in a series expansion and can therefore be controlled at will. It appears that the first few terms are usually adequate. The exact result exhibits a phenomenon that we call asymmetric screening. Namely, the magnitude of attractive interactions is decreased (relative to point charges in an infinite solvent) while the magnitude of repulsive interactions is increased (again, relative to point charges in an infinite solvent). This effect might aid in the adoption of correct conformations and in intermolecular recognition. Evaluation of the energy involves only matrix inversion. The surface charge method can be transformed easily to a numerical method for use with arbitrary surfaces. With modest additions, the model also describes an electrorheological fluid. Such a system provides the cleanest opportunity to apply the model.

DOI: 10.1103/PhysRevE.73.061902 PACS number(s): 87.15.Aa, 83.80.Hj, 83.80.Gv, 41.20.Cv

I. MOTIVATION AND CURRENT METHODS

Electrostatic interactions in biological systems are both important and difficult to calculate accurately in practice. Many biological molecules bear considerable electric charge. For example, DNA typically bears a linear charge density of $\sim -6e/\mathrm{nm}$. Several of the amino acids that form proteins form ions in solution. The solvent itself—namely, water—produces considerable electrostatic effects. Water has a high dielectric constant ($\epsilon \approx 73$ at 40 °C) [1] compared to, for example, proteins ($\epsilon \approx 3-5$) [2] and therefore creates considerable induced charge densities during screening. On top of all this, numerous ions of various physical sizes and with various electric charges inhabit the solvent and are functionally important, as attested to by the existence of the various ion channels ([3] is one of many articles on ion channels). Furthermore, hydrogen bonds, known to be involved in helix formation in both DNA and proteins, are essentially electrostatic in origin. Indeed, it seems that electrostatic effects often drive the physical-chemical processes in biological systems and, thereby, determine biological function. Therefore, any attempt to perform molecular dynamics (MD) simulations of biological systems will require an adequate description of these electrostatic forces.

In order to focus the discussion, consider a protein in water. Considerable interest exists in predicting and understanding the physical basis of the protein's interactions with other molecules and its conformation, as well as the path in conformation space that the protein follows on its way to its native conformation. At first glance, computing the electrostatic interactions described above on the classical level might seem to be a simple matter. After all, Maxwell's equations are known. In practice, however, the task is not so simple, a fact attested to by the rather large body of literature, including several fairly comprehensive review articles [2,4–6], that has been developed over the years, often under the term "solvation."

Because the electrostatic energy needs to be calculated at each step of a MD simulation, MD requires a method that, in addition to being accurate, is computationally fast. There is often a trade-off between these two basic requirements. The methods used to deal with solvation in MD simulations can be divided into explicit and implicit solvent methods. The eplicit solvent methods, as the name indicates, include water molecules (and perhaps ions) in the MD simulation in much the same way as the other molecules are included, although separate effort is often made to parametrize the representation of the solvent. In some cases, treating the solvent molecules on an equal footing with the solute molecule(s) might even seem to be necessary, as, for example, when a small number of solvent molecules are intimately contained within the structure of a larger solute. In such cases, a number of solvent molecules are part and parcel of the solute. The advantage of using explicit solvent is that, to the extent that the force field and its parametrization are valid, the results should reflect reality. However, there are at least two disadvantages to this approach. First, considerable computational effort is expended on molecules of little intrinsic interest. Second, decisions regarding the amoung of water to include, the boundary conditions, and the parametrization of the water will often partially undermine the advantage of explicit solvent. Often, the solvent can be conceptually distinguished from the solute, and in such cases it is generally the solute that is of primary interest, the solvent being of interest only insofar as it has an effect on the solute. In such cases, it is desirable, if possible, to use methods that treat the solvent in an average way. Such methods are called implicit or continuum solvent methods. One possibility is to attempt to average over the atomic polarizations to obtain a position-

*Electronic address: doerr@ncbi.nlm.nih.gov
†Electronic address: yyu@ncbi.nlm.nih.gov

dependent dielectric function [7]. Sometimes referred to as Lorentz-Debye-Sack (LDS) theory, this approach in practice generally produces a distance-dependent (i.e., spherically symmetric) screening function. At times, LDS theory has been combined with aspects of the Born model discussed below [7,8]. Another more frequently used possibility is to regard the dielectric constants of the solvent and solute as parameters of the model. The continuum solvent methods in common use in MD simulations can be divided into two main groups: methods that solve the Poisson equation [or the Poisson-Boltzmann (PB) equation if ions in the solvent are taken into account] and methods that use some version of the generalized Born (GB) model.

The PB methods have the virtue of being straightforward in principle, since the strategy is just to solve the PB equation numermically. Many different schemes have been devised to accomplish this [9]. PB methods are intermediate in computational time with respect to explicit solvent methods and GB methods. Furthermore, because PB methods yield a numerical electrostatic potential, forces have to be computed as numerical gradients.

GB methods are extentions of one kind or another of Born's model of solvation of an ion [10]. In the Born model, an ion in solution is regarded as a spherical cavity of radius a in an infinite dielectric medium (dielectric constant ϵ_o) with a point charge q at the center of the cavity. The exact heat of hydration is $\frac{1}{2}\left(\frac{1}{\epsilon_0}-1\right)\frac{q^2}{a}$. The motivation for the GB approach arises from the observation that the difference in energy between a collection of such ions (charges q_i) with large pairwise separations r_{ij} in an infinite dielectric medium (dielectric constant ϵ_0) and a similar collection in vacuum is $\frac{1}{2}\left(\frac{1}{\epsilon_0}-1\right)\Sigma_{i\neq j}\frac{q_i q_j}{r_{ij}}$. Given the similarity of the two expressions, it is perhaps natural to inquire whether one can find a formula of this general form that has these limits. The most commonly used expression is that of Still *et al.* [11]:

$$\left[r_{ij}^2+\alpha_i\alpha_j\exp\left(-\frac{r_{ij}^2}{4\alpha_i\alpha_j}\right)\right]^{1/2},$$

where α_i is the Born radius of atom i. Although the Born radii are just free parameters in the context of the justification given above for the GB model, a somewhat more formal justification can be given for the values given the Born radii. The derivation involves the Coulomb field approximation (CFA). For the original Born model of a single ion in an infinite solvent, the electric displacement of the point charge is the same as if there were no solvent at all. This is true because of the spherical symmetry of the situation. Without the symmetry there is no reason to expect the same result for the displacement. The CFA consists of using this same electric displacement for each point charge in the joint cavity of all the atoms with all the other point charges set to zero. Most of the numerous variations on the simple GB model involve different methods of approximating (to increase speed) or empirically adjusting (to increase accuracy) the Born radii. In the original GB formulation of Still *et al.* [11], the Born radii were expressed as an integral over the solute volume and were computed by numerical integration on a grid. These Born radii depended on the configuration of the solute, but for the purpose of computing derivatives to obtain the force, Still *et al.* ignored this dependence. Schaefer and Froemmel [12] obtained Born radii by computing integrals of the fields with various nonsystematic approximations. In generalized born/analytical continuum electrostatics (GB/ACE), Schaefer and Karplus [13] represented the molecular volume as a superposition of Gaussian functions. Lee, Salisbury, and Brooks [14] represented the molecular volume (which needs to be integrated over in the typical formulas used to generate the Born radii) as a superposition of quartic exponential functions within a switching function and added an empirical correction term to the CFA. Ghosh, Rapp, and Friesner [15] replaced the integrals over the molecular volume with integrals over the molecular surface; this is the surface GB method and should be more efficient. It is worth noting the basic approximations common to most of the GB variations: first, the GB equation itself is a convenient guess as to the form of the energy; second, the CFA is usually used when obtaining the Born radii. The specific schemes have various other additional approximations, some of them noted above.

Two major avenues are being pursued to improve the GB model. The first type of proposed improvement involves moving beyond a point-charge model of the charge distribution of the solute molecule, obviously an approximation of the actual electronic charge density. The point charge model potentially does not have enough flexibility to adequately reflect the behavior of the system. The methods proposed to address this issue include the addition of off-center charges and the use of multipole expansions to represent the solute charge distribution. Another type of proposed improvement addresses the fact that the solute molecule itself, not just the dielectric medium, is polarizable. Many methods do not reflect this degree of freedom. The methods proposed to address this issue include allowing the charge to fluctuate, inclusion of Drude oscillators, and the inclusion of point polarizabilities (usually dipoles). Several other issues deserve additional attention. Due to the complications they present, ions are not consistently included. Also, the force is needed in MD calculations, and if the potential is what is in hand, a derivative is required. It would be nice to avoid or mitigate the process of numerical differentiation. Finally, the choice of surface is, in general, problematic.

II. MODIFIED APPROACH

Suppose one wishes to consider a biomolecule in solution (generally water). It is standard, and appropriate for classical considerations, to model a molecule as a collection of (possibly overlapping) spheres. If, as is often the case, one wishes to proceed at an atomic level, the spheres represent individual atoms. Coarse-grained methods can be implemented by associating a sphere with a larger chemical group (e.g., an amino acid) and choosing parameters accordingly. We shall focus on the atomistic model, although most of our remarks are applicable to the coarse-grained models, as well. Associated with each sphere is an electric charge. In principle, one can also assign higher-order multipole moments (with respect to some coordinate system). Rarely are mo-

ments higher than the dipole assigned. Each sphere represents a polarizable (dielectric) medium and, in the presence of external fields such as the fields due to other similar spheres, can produce higher-order multipole moments. The full electric field is then the sum of the fields due to all the spheres including the original monopole moments and the fields due to polarization. It is quite possible to treat the water molecules in much the same manner as was just described for the solute molecule. However, as mentioned above, this explicit solvent method is computationally laborious with much labor being spent on water molecules that are not of primary interest. It is tempting to replace the explicit solvent with a continuous dielectric medium everywhere outside the molecule of interest.

The equation that must be solved to find the electric field (or potential) is either Poisson's equation or the Poisson-Boltzmann equation (nonlinear or a linear approximation) depending on whether one is including the effects of ions in the solvent. These differential equations give rise to boundary value problems. The boundaries are at infinity (if a large "containter" is assumed, otherwise a boundary at the container) and at the interface between the spheres and solvent (water). The first instinct when confronted with such a boundary value problem is, if the boundary lacks symmetry or is otherwose sufficiently complex, to solve the differential equation numerically on some appropriate grid. (The usual analytical methods—e.g., separation of variables—are surprisingly involved even for the relatively simple situation of just two dielectric spheres in an infinite medium.) Because the differential equations are second order, there will be two boundary conditions. It should also be noted that such algorithms are implemented on a three-dimensional grid and will therefore generally scale with the volume of the system.

We take a fresh approach. For the time being we will consider Poisson's equation, although the extention to include ions is underway. We model a molecule as a collection of nonoverlapping linear dielectric spheres representing atoms. There is then an interfacial surface between the molecule (collection of spheres) and the solvent. Because the system is a piecewise-constant linear dielectric, the polarization produces induced charge only at the boundary. This well-known fact is a key point, since it points the way to transforming a three-dimensional problem to a two-dimensional problem. The main distinctive feature of our method is the use of the induced surface charge density as the function to be solved for. This strategy has been aluded to earlier [16] and presented in simpler geometries [17]. Other surface methods are difficult to adapt for MD [18,19]. The surface charge strategy is here applied to a system of a large number of spheres; the numerical extension for a general surface is outlined. Given the decision to regard the induced surface charge density on each sphere as the function to be solved for, it is reasonable to expand the induced surface charge density in spherical harmonics. The solution therefore comprises the set of coefficients in these expansions: σ_{lm}^{i}, the l,m component of the surface charge density of the sphere i. Once the potential has been written in terms of the surface charge density, there is only one boundary condition: continuity of the dielectric constant times the normal derivative of the potential at the surface. The potential is automatically continuous at the boundary because it has been expressed directly in terms of the surface charge density there. Once the potential V has been expressed in coordinates suitable for enforcing the boundary condition, one obtains an infinite set of linear algebraic equations for the components of the surface charge density. A solution (possibly approximate) of this set of equations will then allow substitution of these components back into the original expression for the potential.

An advantage of this method is that it transforms a three-dimensional problem into a two-dimensional problem. The benefit of this change is clear for cases in which a numerical solution is necessary. In fact, somewhat similar two-dimensional methods focusing on the surface charge density [sometimes called apparent surface charge (ASC) methods] have been developed [4], but are not in widespread use in MD simulations. They are ocassionally used for one-time calculations of the electric potential of fixed solute molecules [20]. Another advantage of the surface charge method is that the original differential equation has been replaced by a set of linear algebraic equations. In other words, straighforward matrix inversion is the main numerical tool required. The solution presented for a system of spheres is also applicable to a colloidal system [21,22].

III. THE MODEL

Consider N dielectric spheres, $\mathcal{A}_1,\ldots,\mathcal{A}_N$, immersed in an infinite medium of dielectric constant ϵ_0. The ith dielectric sphere $\mathcal{A}_i$ has radius a_i, has a point charge q_i at its center, and consists of a material with dielectric constant ϵ_i. The induced surface charge density on sphere $\mathcal{A}_i$ is σ^i. When considering the contribution to the potential near the surface of $\mathcal{A}_k$ due to the surface charge on $\mathcal{A}_i$, it is convenient to use the following local coordinate system (as opposed to the global laboratory coordinate system). The origin is chosen to be at the center of $\mathcal{A}_k$, and the negative z axis is chosen to pass through the center of $\mathcal{A}_i$. The distance between the centers of $\mathcal{A}_i$ and $\mathcal{A}_k$ is $L_{ik}=|\boldsymbol{R}_i-\boldsymbol{R}_k|$, and the unit vector pointing from the center of $\mathcal{A}_k$ to the center of $\mathcal{A}_i$ is $\boldsymbol{n}_{ik}=(\boldsymbol{R}_i-\boldsymbol{R}_k)/L_{ik}$. (Often, when there is no chance of confusion, the subscripts will be dropped from symbols such as L_{ik} and $\boldsymbol{n}_{ik}$.) The rotation of the laboratory z axis to the local z axis is shown in Fig. 1. The two angles ϑ and φ suffice because there is no preferred orientation for the local x and y axes. The angles ϑ and φ, of course, differ for each local coordinate system.[1] The vector from the center of $\mathcal{A}_l$ to an arbitrary point P is $\boldsymbol{r}_l$. The vector from the center of $\mathcal{A}_l$ to an arbitrary point P on the surface of the sphere $\mathcal{A}_l$ is $\boldsymbol{r}_l'$. The angle between $\boldsymbol{r}_l$ and the local positive z axis is $\tilde{\theta}_l$; the azimuthal angle of $\boldsymbol{r}_l$ with respect to the local coordinate system is $\tilde{\phi}_l$. The angle between $\boldsymbol{r}_l'$ and the local positive z axis is $\tilde{\theta}_l'$; the azimuthal

[1] In the degenerate case that the global and local z axes are parallel, the global and local coordinate systems are chosen to be identical (i.e., $\vartheta=\pi$ and $\varphi=\pi$). In the other degenerate case that the global and local z axes are antiparallel, the new coordinate system is reached by rotating counterclockwise around the global y axis by π (i.e., $\vartheta=0$ and $\varphi=\pi$).

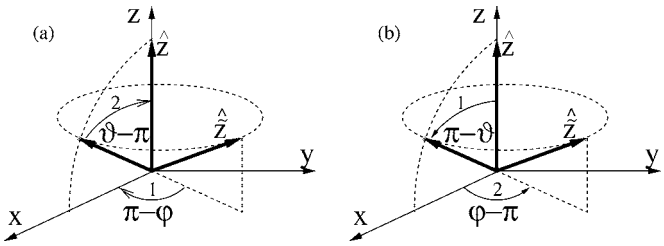

FIG. 1. (a) The rotation that takes the local z axis to the global z axis. A rotation (labeled 1) about the global z axis through an angle $\pi-\varphi$ is followed by a rotation (labeled 2) about the global y axis through an angle $\vartheta-\pi$. In the text, this rotation is denoted by R. (b) The rotation that takes the global z axis to the local z axis. A rotation (labeled 1) about the global y axis through an angle $\pi-\vartheta$ is followed by a rotation (labeled 2) about the global z axis through an angle $\varphi-\pi$. In the text, this rotation is denoted by R^{-1}. The notation for the rotation angles has been chosen for later convenience.

angle of $\boldsymbol{r}_l'$ is $\tilde{\phi}_l'$. The vector to an arbitrary point from the origin is $\boldsymbol{r}$ (no subscript) and has the usual polar coordinates r, θ, and ϕ. The vector to an arbitrary source point (i.e., an arbitrary point on the surface of one of the spheres) from the origin is $\boldsymbol{r}'$ (no subscript) and has the usual polar coordinates r', θ', and ϕ'. The pattern of the notation is that a vector or angle without a prime refers to the location of a general point, a vector or an angle with a prime refers to the location of a source point, an angle without a tilde is with respect to the laboratory coordinate system, and an angle with a tilde is with respect to the local coordinate system. The vector to the center of $\mathcal{A}_l$ from the origin is $\boldsymbol{R}_l$ and has the usual polar coordinates R_l, Θ_l, and Φ_l.

The ultimate objective is to find the total electrostatic energy of such a system of N dielectric spheres embedded in an unbounded dielectric medium. To that end, one wishes to calculate for an arbitrary point in space the electrical potential, which is a linear superposition of the potentials of the screened point charges q_i/ϵ_i $(i=1,\ldots,N)$ and the surface charge densities σ^i $(i=1,\ldots,N)$:

$$V(\boldsymbol{r})=\sum_{i=1}^{N}\frac{q_i}{\epsilon_i|\boldsymbol{r}-\boldsymbol{R}_i|}+\sum_{i=1}^{N}\int_{\mathcal{A}_i}\frac{\sigma^i(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}S_i'. \tag{1}$$

As mentioned in the preceding section, the induced surface charge densities are not known; rather, they are obtained by enforcing the boundary condition at the surface of each sphere. Application of the boundary condition requires that the total potential of the known point charges and the still unspecified surface charge densities be calculated just inside and just outside the surface of each sphere. The boundary conditions will yield algebraic equations which, when solved, give the induced surface charge densities in terms of known quantities. As mentioned above, once the surface charge densities are know, the potential is obtained simply by feeding the expression for the surface charge density back into Eq. (1), the original general expression for the electrostatic potential.

Clearly, the model only generates the electrostatic interactions in a biomolecular system. In order to apply the model in such a context, the result will have to be integrated into a program (MD, for example) that computes the behavior of such a system. However, with modest additions the model also describes an electrorheological fluid. Such a system will offer the cleanest opportunity to apply the model. This application will be presented in a subsequent publication.

IV. TWO SPHERES: AN IMPORTANT SUBSYSTEM

As explained in the preceding section, the central task is to apply the boundary condition at the surface of each sphere. Since each boundary is spherical, it is obviously desirable, for the purpose of applying the boundary condition, to express the potential near each spherical surface in terms of spherical polar coordinates centered on that sphere. For example, consider the potential near the surface $\mathcal{A}_k$. The contributions from the point charges and from the induced surface charge on $\mathcal{A}_k$ are relatively straightforward to calculate in these coordinates. However, the contribution to the potential near $\mathcal{A}_k$ from the induced surface charge on another sphere $\mathcal{A}_i$ requires a little work to obtain. This contribution will now be calculated.

The potential due to the (as yet undetermined) induced surface charge on sphere $\mathcal{A}_i$ is

$$V_i(\boldsymbol{r})=\int_{\mathcal{A}_i}\frac{\sigma^i(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}S_i'. \tag{2}$$

Since our interest lies in the potential due to σ^i near some other sphere $\mathcal{A}_k$, we note the geometrically evident expressions (see Fig. 2) involving $\boldsymbol{r}$ and $\boldsymbol{r}'$: $\boldsymbol{r}=\boldsymbol{R}_k+\boldsymbol{r}_k$ and $\boldsymbol{r}'=\boldsymbol{R}_k+L\boldsymbol{n}+\boldsymbol{r}_i'$. These expressions lead to $\boldsymbol{r}-\boldsymbol{r}'=\boldsymbol{r}_k-(L\boldsymbol{n}+\boldsymbol{r}_i')$, which, with the observation that $r_k\leqslant|L\boldsymbol{n}+\boldsymbol{r}_i'|$ when $r_k\to a_k$, permits the denominator of Eq. (2) to be expanded:

$$|\boldsymbol{r}-\boldsymbol{r}'|^{-1}=|\boldsymbol{r}_k-(L\boldsymbol{n}+\boldsymbol{r}_i')|^{-1}=\sum_{l'=0}^{\infty}r_k^{l'}|L\boldsymbol{n}+\boldsymbol{r}_i'|^{-l'-1}P_{l'}(\cos\delta),$$

where δ is the angle between $\boldsymbol{r}_k$ and $L\boldsymbol{n}+\boldsymbol{r}_i'$. Furthermore,

$$P_{l'}(\cos\delta)=\frac{4\pi}{2l'+1}\sum_{m'=-l'}^{l'}Y_{l'm'}(\tilde{\theta}_k,\tilde{\phi}_k)Y_{l'm'}^{*}(\hat{\theta}',\hat{\phi}'),$$

where $\hat{\theta}'$ and $\hat{\phi}'$ are the spherical polar angles of the vector $L\boldsymbol{n}+\boldsymbol{r}_i'$ with respect to the local coordinate system with origin at the center of $\mathcal{A}_k$. The surface charge density can also be expanded in spherical harmonics:

$$\begin{aligned}\sigma^i(\boldsymbol{r}')&=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sqrt{4\pi}\sigma_{lm}^iY_{lm}(\theta_i',\phi_i')\\&=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sqrt{4\pi}\tilde{\sigma}_{lm}^iY_{lm}(\tilde{\theta}_i',\tilde{\phi}_i').\end{aligned} \tag{3}$$

This expansion can be made with to respect to any set of coordinate axes one may care to choose. In the present instance, it is most convenient to use the expansion with respect to the local coordinate system [the second expansion in Eq. (3)] in order to carry out the integration in Eq. (2). How-

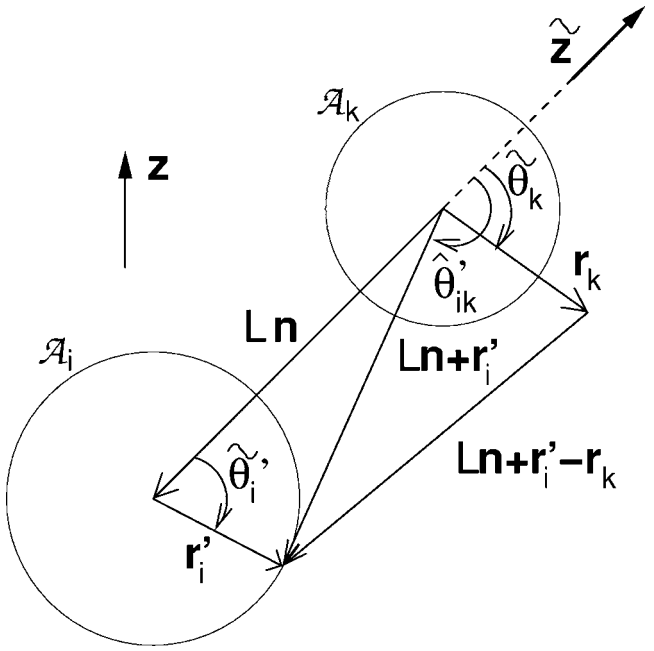

FIG. 2. The geometry of the two-sphere subsystem (spheres $\mathcal{A}_k$ and $\mathcal{A}_i$) with some of the useful variables indicated: $\boldsymbol{r}_k$ is a vector from the center of $\mathcal{A}_k$ to an arbitrary point, $\boldsymbol{r}_i'$ is a vector from the center of $\mathcal{A}_i$ to a source on the surface of $\mathcal{A}_i$, $L\boldsymbol{n}$ is a vector from the center of $\mathcal{A}_k$ to the center of $\mathcal{A}_i$, $\tilde{\theta}_i'$ is the polar angle of $\boldsymbol{r}_i'$ in the local coordinate system, $\tilde{\theta}_k$ is the polar angle of $\boldsymbol{r}_k$ in the local coordinate system, $\hat{\theta}_{ik}'$ is the polar angle of $L\boldsymbol{n}+\boldsymbol{r}_i'$ in the local coordinate system, $\boldsymbol{z}$ is a unit vector parallel to the global z axis, and $\tilde{\boldsymbol{z}}$ is a unit vector parallel to the local z axis. Two other vectors, derived from the others and prominent in the analysis, are also shown.

ever, it will ultimately be necessary to transform back to global coordinates in order to apply the boundary condition. The two expansions are, of course, not independent; the connection between σ_{lm}^i and $\tilde{\sigma}_{lm}^i$ is presented in the Appendix. Upon using the expansions for the numerator and the denominator in Eq. (2), the potential becomes

$$V_i(\boldsymbol{r}_k) = \sum_{lml'm'} \frac{\sqrt{4\pi}\tilde{Q}_{lm}^i}{2l'+1} Y_{l'm'}(\tilde{\theta}_k, \tilde{\phi}_k) r_k^{l'} \times \int \mathrm{d}(\cos\tilde{\theta}_i')\mathrm{d}\tilde{\phi}_i' \frac{Y_{lm}(\tilde{\theta}_i', \tilde{\phi}_i') Y_{l'm'}^*(\hat{\theta}', \hat{\phi}')}{|L\boldsymbol{n}+\boldsymbol{r}_i'|^{l'+1}}, \quad (4)$$

where $\tilde{Q}_{lm}^i = 4\pi a_i^2 \tilde{\sigma}_{lm}^i$. [Since $\tilde{Q}_{lm}^i$ depends on which of the many local coordinate systems the surface charge is being represented in, $\tilde{Q}_{lm}^i$ is actually a simplified notation for $\tilde{Q}_{lm}^i(k)$.]

In order to carry out the angular integrations in Eq. (4), $\hat{\theta}'$, $\hat{\phi}'$, and $|L\boldsymbol{n}+\boldsymbol{r}_i'|$ must be expressed in terms of the integration variables $\tilde{\theta}_i'$ and $\tilde{\phi}_i'$. The law of cosines gives $|L\boldsymbol{n}+\boldsymbol{r}_i'| = L(1+t_{ik}^2 - 2t_{ik}\cos\tilde{\theta}_i')^{1/2}$ where $t_{ik} \equiv a_i/L_{ik}$ (for brevity, t_{ik} will be shortened to t_i when this will not cause confusion). For the azimuthal angle one may choose $\hat{\phi}' = \tilde{\phi}_i'$. Finally, geometrical considerations (see Fig. 2) give $|L\boldsymbol{n}+\boldsymbol{r}_i'|\cos\hat{\theta}' = a_i\cos\tilde{\theta}_i' - L$ which leads to

$$\cos\hat{\theta}' = \frac{t_i\cos\tilde{\theta}_i' - 1}{(1+t_i^2 - 2t_i\cos\tilde{\theta}_i')^{1/2}}.$$

Now Eq. (4) becomes

$$V_i(\boldsymbol{r}_k) = \sum_{lml'} \frac{\sqrt{4\pi}\tilde{Q}_{lm}^i r_k^{l'}}{2(2l'+1)L^{l'+1}} Y_{l'm}(\tilde{\theta}_k, \tilde{\phi}_k) \times \sqrt{\frac{(2l+1)(l-m)!\,(2l'+1)(l'-m)!}{(l+m)!\,(l'+m)!}} \times \int_{-1}^{1} \frac{P_{lm}(x) P_{l'm}\left(\frac{t_i x - 1}{\sqrt{1+t_i^2-2t_i x}}\right)}{(1+t_i^2-2t_i x)^{(l'+1)/2}} \mathrm{d}x,$$

where $x = \cos\tilde{\theta}_i'$. The integral has been motivated and evaluated by Yu [16]:

$$\int_{-1}^{1} \frac{P_{lm}(x) P_{l'm}\left(\frac{tx-1}{\sqrt{1+t^2-2tx}}\right)}{(1+t^2-2tx)^{(l'+1)/2}} \mathrm{d}x = \frac{2t^l(-1)^{l'-m}(l+l')!}{(2l+1)(l-m)!\,(l'-m)!}.$$

The potential becomes

$$V_i(\boldsymbol{r}_k) = \sum_{lml'} \frac{\sqrt{4\pi}\tilde{Q}_{lm}^i t_i^l (-1)^{l'-m}(l+l')!}{L^{l'+1}\sqrt{(2l+1)(2l'+1)(l+m)!\,(l'+m)!\,(l-m)!\,(l'-m)!}} r_k^{l'} Y_{l'm}(\tilde{\theta}_k, \tilde{\phi}_k). \quad (5)$$

Finally, in anticipation of enforcing the boundary condition in the global coordinate system, we make (based upon the analysis in the Appendix) the following substitutions:

$$Y_{l'm}(\tilde{\theta}_k,\tilde{\phi}_k)=\langle\tilde{\theta}_k,\tilde{\phi}_k|l',m\rangle=\langle\theta,\phi|\mathcal{D}^{\dagger}(R)|l',m\rangle=\sum_{m''}\mathcal{D}^{(l')}_{m''m}(R^{-1})Y_{l'm''}(\theta_k,\phi_k),$$

$$\tilde{Q}^i_{lm}=\sum_{m'}\mathcal{D}^{(l)}_{mm'}(R)Q^i_{lm'}.$$

The rotation R, which is actually $R_{i\to k}$ with suppressed indices, takes the local z axis, which points from the center of sphere i to the center of sphere k, to the global z axis. Along with a relabeling of summation indices ($l\leftrightarrow l'$ and $m\leftrightarrow m''$), these substitutions transform Eq. (5) into

$$V_i(\boldsymbol{r}_k)=\sum_{lmm'm''l'}\frac{\sqrt{4\pi}\mathcal{D}^{(l')}_{m''m'}(R)Q^i_{l'm'}t^{l'}_i(-1)^{l-m''}(l+l')!}{L^{l+1}\sqrt{(2l+1)(2l'+1)(l+m'')!\,(l'+m'')!\,(l-m'')!\,(l'-m'')!}}r^l_k\mathcal{D}^{(l)}_{mm''}(R^{-1})Y_{lm}(\theta_k,\phi_k).\tag{6}$$

V. MANY SPHERES

The previous section was devoted to obtaining an expression for the potential near $\mathcal{A}_k$ due to the surface charge on $\mathcal{A}_i$ ($i\neq k$) written in terms of spherical coordinates centered on $\mathcal{A}_k$, with an eye toward enforcing the boundary condition at $\mathcal{A}_k$. The others terms in Eq. (1) for the electrostatic potential will now be obtained.

First, consider the potential near $\mathcal{A}_k$ due to the surface charge on $\mathcal{A}_k$ itself:

$$V_k(\boldsymbol{r}_k)=\int_{\mathcal{A}_k}\frac{\sigma^k(\boldsymbol{r}'_k)}{|\boldsymbol{r}_k-\boldsymbol{r}'_k|}\mathrm{d}S'_k.$$

The numerator may be expanded as before, except that it is not necessary to switch to a rotated coordinate system:

$$\sigma^k(\boldsymbol{r}'_k)=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sqrt{4\pi}\sigma^k_{lm}Y_{lm}(\theta'_k,\phi'_k).$$

The expansion of the denominator is

$$|\boldsymbol{r}_k-\boldsymbol{r}'_k|^{-1}=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\frac{4\pi}{2l+1}r^l_<r_>^{-l-1}Y^*_{lm}(\theta'_k,\phi'_k)Y_{lm}(\theta_k,\phi_k),$$

where $r_<=\min(r_k,a_k)$ and $r_>=\max(r_k,a_k)$. One finds

$$V_k(\boldsymbol{r}_k)=\sum_{lm}\frac{\sqrt{4\pi}Q^k_{lm}r^l_<}{(2l+1)r_>^{l+1}}Y_{lm}(\theta_k,\phi_k),\tag{7}$$

with $Q^k_{lm}=4\pi a_k^2\sigma^k_{lm}$.

Second, consider the potential of q_i near $\mathcal{A}_k$ ($i\neq k$). The dielectric medium in which q_i is placed causes the charge to be screened to q_i/ϵ_i. Therefore, because $r_k<L$, the potential is

$$\frac{q_i}{\epsilon_i|\boldsymbol{r}_k-L\boldsymbol{n}|}=\frac{q_i}{\epsilon_i}\sum_{l=0}^{\infty}\frac{r^l_k}{L^{l+1}}\frac{4\pi}{2l+1}\sum_{m=-l}^{l}Y_{lm}(\theta_k,\phi_k)Y^*_{lm}(\vartheta,\varphi).\tag{8}$$

The angles ϑ and φ are the polar angles of $L\boldsymbol{n}$ in the global reference frame (see Figs. 1 and 2).

Finally, the potential of q_k near $\mathcal{A}_k$ is just

$$\frac{q_k}{\epsilon_k r_k}\sqrt{4\pi}Y_{00}(\theta_k,\phi_k).\tag{9}$$

The full potential in the vicinity of $\mathcal{A}_k$ is then found by substitution of Eqs. (6)–(9) into Eq. (1). With the full potential written in convenient coordinates, one can now enforce the boundary condition at $\mathcal{A}_k$ for each k:

$$\epsilon_k\left.\frac{\partial V^{\mathrm{in}}}{\partial r_k}\right|_{a_k}=\epsilon_0\left.\frac{\partial V^{\mathrm{ex}}}{\partial r_k}\right|_{a_k}.$$

For the case $l=m=0$, the boundary condition reduces to a previously known fact [17] about the total surface charge on $\mathcal{A}_k$:

$$Q^k_{00}=Q^k=q_k\left(\frac{1}{\epsilon_0}-\frac{1}{\epsilon_k}\right).\tag{10}$$

For all other terms ($l>0$), one finds

$$\frac{[\epsilon_k l+\epsilon_0(l+1)]\sqrt{4\pi}}{(2l+1)}Q^k_{lm}=(\epsilon_0-\epsilon_k)\left[\sum_{i\neq k}\frac{q_i l t^{l+1}_{ki}4\pi}{\epsilon_i(2l+1)}Y^*_{lm}(\vartheta_{ik},\varphi_{ik})+\sum_{\substack{l',m',m''\\ i\neq k}}\frac{\sqrt{4\pi}\mathcal{D}^{(l')}_{m''m'}(R_{i\to k})Q^i_{l'm'}t^{l'}_{ik}(-1)^{l-m''}(l+l')!\,l t^{l+1}_{ki}\mathcal{D}^{(l)}_{mm''}(R^{-1}_{i\to k})}{[(2l+1)(2l'+1)(l-m'')!\,(l'-m'')!\,(l+m'')!\,(l'+m'')!\,]^{1/2}}\right],\tag{11}$$

which, in principle, is to be solved for the Q^i_{lm} for all i, $l>0$, and $-l\leqslant m\leqslant l$. For clarity, the dependence of $t_{ki}=a_k/L_{ik}$ on its second index has been restored; the indices of the angles ϑ_{ik} and φ_{ik} have also been restored, as have the indices of $R_{i\to k}$.

VI. ELECTROSTATIC ENERGY

The electrostatic energy U can be obtained from the potential

$$U=\sum_{i=1}^{N}\frac{1}{2}q_i\left[\sum_{j\neq i}\frac{q_j}{\epsilon_j|\boldsymbol{R}_i-\boldsymbol{R}_j|}+\sum_{j=1}^{N}V_j(\boldsymbol{r}=\boldsymbol{R}_i)\right]$$
$$=\sum_{i=1}^{N}\sum_{j\neq i}\left[\frac{q_iq_j}{2\epsilon_jL_{ij}}+\frac{1}{2}q_iV_j(\boldsymbol{r}_i=\boldsymbol{0})\right]+\sum_{i=1}^{N}\frac{1}{2}q_iV_i(\boldsymbol{r}_i=\boldsymbol{0}). \tag{12}$$

In each case, the potentials are evaluated at the location of the charge q_i. The final term in Eq. (12) involves Eq. (7) with $\boldsymbol{r}_i=\boldsymbol{0}$. But $\boldsymbol{r}_i=\boldsymbol{0}$ implies that $r_<=0$ and $r_>=a_i$, so only the single term with $l=m=0$ survives in Eq. (7). Using Eq. (10), one finds, for the final term in Eq. (12),

$$\sum_{i=1}^{N}\frac{q_i^2}{2a_i}\left(\frac{1}{\epsilon_0}-\frac{1}{\epsilon_i}\right).$$

Now consider the first part of Eq. (12). This term involves the potential

$$V_j(\boldsymbol{r}_i=\boldsymbol{0})=\sum_{l=0}^{\infty}\frac{\tilde{Q}^j_{l0}(i)t^l_{ji}}{L_{ij}\sqrt{2l+1}}=\frac{\tilde{Q}^j_{00}}{L_{ij}}+\sum_{l=1}^{\infty}\frac{\tilde{Q}^j_{l0}(i)t^l_{ji}}{L_{ij}\sqrt{2l+1}}$$
$$=\frac{Q^j_{00}}{L_{ij}}+\sum_{l=1}^{\infty}\frac{\tilde{Q}^j_{l0}(i)t^l_{ji}}{L_{ij}\sqrt{2l+1}},$$

where the i dependence of $\tilde{Q}^j_{l0}(i)$ due to the coordinate change has been made explicit. With this expression, the first term in Eq. (12) becomes

$$\sum_{i=1}^{N}\sum_{j\neq i}\frac{q_iq_j}{2\epsilon_0L_{ij}}+\sum_{i=1}^{N}\sum_{j\neq i}\sum_{l=1}^{\infty}\frac{q_i\tilde{Q}^j_{l0}(i)t^l_{ji}}{2L_{ij}\sqrt{2l+1}}.$$

One has, for the energy,

$$U=\sum_{i=1}^{N}\frac{q_i^2}{2a_i}\left(\frac{1}{\epsilon_0}-\frac{1}{\epsilon_i}\right)+\sum_{i=1}^{N}\sum_{j\neq i}\frac{q_iq_j}{2\epsilon_0L_{ij}}+\sum_{i=1}^{N}\sum_{j\neq i}\sum_{l=1}^{\infty}\frac{q_i\tilde{Q}^j_{l0}(i)t^l_{ji}}{2L_{ij}\sqrt{2l+1}}. \tag{13}$$

The first term is the so-called Born self-energy and represents the energy difference upon transferring each of the charged dielectric spheres from vacuum to the medium of dielectric constant ϵ_0. The second sum is the lowest order of interaction between each two spheres. In particular, as two spheres are separated by ever larger distances, the energy becomes simply the Coulomb interaction of the two point charges embedded in the medium of dielectric constant ϵ_0. The final multiple sum represents the corrections to the previous two contributions. Notice that each element of the sum is proportional to one of the point charges and to one of the spherical harmonic components of the induced surface charge density on one of the other spheres.

VII. SPECIAL CASE: TWO ISOLATED SPHERES

Because some of the existing approximate methods for dealing with the electrostatic interactions under consideration treat each pairwise interaction as occuring in isolation, it is useful to examine the special case of two isolated spheres.

A. Simplification of the linear system of equations

Specifying the spheres with the subscripts (or superscripts) a and b, we use the following notation for this special case: $a_a\to a$, $a_b\to b$, $L_{ab}\to L$. Furthermore, without loss of generality, $\boldsymbol{R}_a=\boldsymbol{0}$ and $\boldsymbol{R}_b=L\hat{z}$. The result of Eq. (10) for the case of two isolated spheres is

$$Q^a_{00}=Q^a=q_a\left(\frac{1}{\epsilon_0}-\frac{1}{\epsilon_a}\right),$$

$$Q^b_{00}=Q^b=q_b\left(\frac{1}{\epsilon_0}-\frac{1}{\epsilon_b}\right). \tag{14}$$

The values of Q^a_{lm} and Q^b_{lm} for $l>0$ are determined by Eq. (11) with only the two spheres a and b:

$$\frac{[\epsilon_bl+\epsilon_0(l+1)]\sqrt{4\pi}}{(2l+1)b^2}Q^b_{lm}$$
$$=(\epsilon_0-\epsilon_b)\left[\frac{q_a}{\epsilon_a}\frac{lb^{l-1}}{L^{l+1}}\frac{4\pi}{2l+1}Y^*_{lm}(\pi,\varphi)+\sum_{l',m',m''}\frac{\sqrt{4\pi}\mathcal{D}^{(l')}_{m''m'}(I)Q^a_{l'm'}t^{l'}_a(-1)^{l-m''}(l+l')!\,lb^{l-1}\mathcal{D}^{(l)}_{mm''}(I)}{L^{l+1}[(2l+1)(2l'+1)(l-m'')!\,(l'-m'')!\,(l+m'')!\,(l'+m'')!\,]^{1/2}}\right],$$

$$\frac{[\epsilon_a l+\epsilon_0(l+1)]\sqrt{4\pi}}{(2l+1)a^2}Q^a_{lm}=(\epsilon_0-\epsilon_a)\left[\frac{q_b}{\epsilon_b}\frac{la^{l-1}}{L^{l+1}}\frac{4\pi}{2l+1}Y^*_{lm}(0,\varphi)\right.$$
$$\left.+\sum_{l',m',m''}\frac{\sqrt{4\pi}\mathcal{D}^{(l')}_{m''m'}(I')Q^b_{l'm'}t_b^{l'}(-1)^{l-m''}(l+l')!\,la^{l-1}\mathcal{D}^{(l)}_{mm''}(I')}{L^{l+1}[(2l+1)(2l'+1)(l-m'')!\,(l'-m'')!\,(l+m'')!\,(l'+m'')!]^{1/2}}\right]. \quad (15)$$

Note that $Y^*_{lm}(\pi,\varphi)=\delta_{m0}P_l(-1)\sqrt{\frac{2l+1}{4\pi}}=\delta_{m0}(-1)^l\sqrt{\frac{2l+1}{4\pi}}$ and $Y^*_{lm}(0,\varphi)=\delta_{m0}P_l(1)\sqrt{\frac{2l+1}{4\pi}}=\delta_{m0}\sqrt{\frac{2l+1}{4\pi}}$. The identity rotation is denoted by I, and the rotation that reverses the direction of the z axis is denoted by I'. The relevant rotation matrix elements are $D^l_{m'm}(I)=\delta_{m'm}$ and $D^l_{m'm}(I')=\delta_{m,-m'}(-1)^{l+m}$. On symmetry grounds we expect a nontrivial result for Q^a_{l0} and Q^b_{l0}, but only the trivial solution $Q^a_{lm}=Q^b_{lm}=0$ for $m\neq 0$. Notice that different values of m decouple. For each value of m, all values of $l\geq|m|$ for the same m couple.

First, we shall check that only the trivial solution exists for Eq. (15) in the case that $m\neq 0$. In this case, the first term on the right-hand side of each equation in Eq. (15) vanishes and the equations become linear homogeneous algebraic equations, albeit an infinite set of such equations:

$$Q^b_{lm}=\sum_{l'=|m|}^{\infty}\frac{(1-x_b)(2l+1)t_a^{l'}(-1)^{l-m}(l+l')!\,lt_b^{l+1}}{[(x_b+1)l+1]\sqrt{(2l+1)(2l'+1)(l-m)!\,(l'-m)!\,(l+m)!\,(l'+m)!}}Q^a_{l'm}\equiv\sum_{l'=|m|}^{\infty}B_{ll'}Q^a_{l'm},$$

$$Q^a_{lm}=\sum_{l'=|m|}^{\infty}\frac{(1-x_a)(2l+1)t_b^{l'}(-1)^{l'+m}(l+l')!\,lt_a^{l+1}}{[(x_b+1)l+1]\sqrt{(2l+1)(2l'+1)(l-m)!\,(l'-m)!\,(l+m)!\,(l'+m)!}}Q^b_{l'm}\equiv\sum_{l'=|m|}^{\infty}A_{ll'}Q^b_{l'm},$$

where A and B are matrices defined by these equations, and we have defined $x_a=\epsilon_a/\epsilon_0$ and $x_b=\epsilon_b/\epsilon_0$. If follows that $\boldsymbol{Q}^a=A\boldsymbol{Q}^b=A(B\boldsymbol{Q}^a)=AB\boldsymbol{Q}^a$, which can only be true either if $\boldsymbol{Q}^a=\boldsymbol{Q}^b=\boldsymbol{0}$ (i.e., the trivial solution) or if $AB=BA=I$. If it can be demonstrated that $AB\neq I$, then only the trivial solution exists. A representative matrix element of BA is

$$[BA]_{ll''}=\frac{(1-x_a)(1-x_b)(2l+1)(-1)^{l+l''}lt_b^{l+l''+1}t_a}{[(x_b+1)l+1]\sqrt{(2l+1)(2l''+1)(l-m)!\,(l+m)!\,(l''+m)!\,(l''-m)!}}$$
$$\times\sum_{l'=|m|}^{\infty}\frac{(t_a^2)^{l'}l'(l+l')!\,(l'+l'')!}{[(x_a+1)l'+1](l'-m)!\,(l'+m)!}$$
$$=\frac{(1-x_a)(1-x_b)(2l+1)(-1)^{l+l''}lt_b^{l+l''+1}t_a}{[(x_b+1)l+1]\sqrt{(2l+1)(2l''+1)(l-m)!\,(l+m)!\,(l''+m)!\,(l''-m)!}}$$
$$\times\frac{(t_a^2)^m(l+m)!\,(l''+m)!\,m}{(x_a+1)(2m)![m+(x_a+1)^{-1}]}$$
$$\times{}_4F_3\left(m+1,l+m+1,l''+m+1,m+\frac{1}{x_a+1};2m+1,m+\frac{x_a+2}{x_a+1},m;t_a^2\right),$$

where ${}_4F_3$ is a hypergeometric function [23]. Make the definition

$$\Lambda_{ll''}=\frac{(1-x_a)(1-x_b)(2l+1)(-1)^{l+l''}lt_b^{l+l''+1}t_a}{[(x_b+1)l+1]\sqrt{(2l+1)(2l''+1)(l-m)!\,(l+m)!\,(l''+m)!\,(l''-m)!}}$$

and then consider the following function of a complex variable s:

$$F_l(s) = \sum_{l'=|m|}^{\infty} \frac{(t_a^2)^{l'} l'(l+l')!\,(l'+l)!}{[(x_a+1)l'+1](l'-m)!\,(l'+m)!} \left(\frac{1}{1-s}\right)^{l+l'+1}, \tag{16}$$

which is analytic everywhere except at $s=1$ where there is a pole of order $l+l'+1 \geqslant 3$. A diagonal element is obtained by evaluating F_l at $s=0$: $\Lambda_{ll}F_l(0)=[BA]_{ll}$. The derivative of F_l evaluated at $s=0$ results in the neighboring off-diagonal element $\Lambda_{ll+1}F_l'(0)=[BA]_{ll+1}$. In general, the nth derivative of F_l evaluated at $s=0$ gives an element in the lth row of BA: $\Lambda_{ll+n}F_l^{(n)}(0)=[BA]_{ll+n}$. Evidently, F_l is a generating function for a sequence of elements in the lth row of BA. If a nontrivial solution is to exist, then we must have $\Lambda_{ll+n}F_l^{(n)}(0) = \delta_{n,0}$. Assuming that this is so leads to a contradiction. By the assumption, $\Lambda_{ll}F_l(0)=1$—i.e., $F_l(0)\neq 0$—and all the derivatives of F_l vanish at $s=0$. But F_l is analytic in the vicinity of $s=0$, and so its power series clearly indicates that F_l is a positive constant η—namely, $F_l(s)=\eta\neq 0$—in the vicinity of $s=0$, and therefore by analytic continuation $F_l(s)=\eta$ everywhere except $s=1$. But as $s\to-\infty$, $F_l(s)\to 0$, which is a contradiction. Therefore, as expected, Eq. (15) has only the trivial solution, and $Q_{lm}^a=Q_{lm}^b=0$ for $m\neq 0$.

A physical interpretation can be given for the s-dependent factor in Eq. (16). For $l=l''$, this factor essentially rescales t_a and t_b: $t_a\to t_a/\sqrt{1-s}$ and $t_b\to t_b/\sqrt{1-s}$. If the identity matrix is to be obtained, we expect $\Lambda_{ll}F_l$ to be invariant under this rescaling. However, as $s\to-\infty$, the two spheres become infinitely separated and the point charge result is regained. In particular, the surface charges vanish. By the expected invariance, we then obtain an unphysical result: the surface charges vanish for all separations. This reasoning does not hold for $m=0$ since those components do not obey the homogeneous equation under consideration here.

Only Q_{l0}^a and Q_{l0}^b remain to be determined. The following definitions will be more natural for the azimuthally symmetric special case now being delt with because they render the expansion for the surface charge densities purely in terms of Legendre polynomials:

$$Q_l^a = \sqrt{2l+1}\,Q_{l0}^a,$$

$$Q_l^b = \sqrt{2l+1}\,Q_{l0}^b.$$

From Eq. (15), we find the equations for Q_l^a and Q_l^b:

$$Q_l^b = \frac{q_a(1-x_b)l(2l+1)(-1)^l t_b^{l+1}}{\epsilon_0[(x_b+1)l+1]x_a} + \frac{(1-x_b)l(2l+1)(-1)^l t_b^{l+1}}{[(x_b+1)l+1]l!} \sum_{l'=0}^{\infty} \frac{(l+l')!\,t_a^{l'} Q_{l'}^a}{l'!\,(2l'+1)},$$

$$Q_l^a = \frac{q_b(1-x_a)l(2l+1)t_a^{l+1}}{\epsilon_0[(x_a+1)l+1]x_b} + \frac{(1-x_a)l(2l+1)t_a^{l+1}}{[(x_a+1)l+1]l!} \sum_{l'=0}^{\infty} \frac{(l+l')!\,t_b^{l'} Q_{l'}^b(-1)^{l'}}{l'!\,(2l'+1)}. \tag{17}$$

A bit more notation will simplify the appearance of these equations. Make the following definitions:

$$\mathcal{Q}_l^b = \frac{(-1)^l Q_l^b t_b^l}{2l+1},$$

$$\mathcal{Q}_l^a = \frac{Q_l^a t_a^l}{2l+1},$$

$$\alpha_l^b = \frac{(1-x_b)l t_b^{2l+1}}{(x_b+1)l+1},$$

$$\alpha_l^a = \frac{(1-x_a)l t_a^{2l+1}}{(x_a+1)l+1},$$

$$\beta_{l,l'} = \frac{(l+l')!}{l!\,l'!} = \beta_{l',l},$$

$$\beta_{l,l'}^a = \alpha_l^a \beta_{l,l'},$$

$$\beta_{l,l'}^b = \alpha_l^b \beta_{l,l'}.$$

Once the first term on the left-hand side of each member of Eq. (17) is combined with the $l'=0$ term from the corresponding sum, Eq. (17) can be cast conveniently in matrix form

$$\underbrace{\left(\begin{array}{ccc|ccc} 1 & 0 & \cdots & -\beta_{11}^a & -\beta_{12}^a & \cdots \\ 0 & 1 & \cdots & -\beta_{21}^a & -\beta_{22}^a & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \hline -\beta_{11}^b & -\beta_{12}^b & \cdots & 1 & 0 & \cdots \\ -\beta_{21}^b & -\beta_{22}^b & \cdots & 0 & 1 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{array}\right)}_{T} \underbrace{\left(\begin{array}{c} \mathcal{Q}_1^a \\ \mathcal{Q}_2^a \\ \vdots \\ \mathcal{Q}_1^b \\ \mathcal{Q}_2^b \\ \vdots \end{array}\right)}_{\mathcal{Q}} = \underbrace{\left(\begin{array}{c} \frac{q_b\alpha_1^a}{\epsilon_0} \\ \frac{q_b\alpha_2^a}{\epsilon_0} \\ \vdots \\ \frac{q_a\alpha_1^b}{\epsilon_0} \\ \frac{q_a\alpha_2^b}{\epsilon_0} \\ \vdots \end{array}\right)}_{X}. \tag{18}$$

Because $T\mathcal{Q}=X$, $\mathcal{Q}$ is given by

$$\mathcal{Q} = T^{-1}X. \tag{19}$$

B. Electrostatic energy

The electrostatic energy for this special case of two isolated spheres follows from Eq. (13) and the additional notation that has been defined:

$$U = \frac{q_a q_b}{\epsilon_0 L} + \frac{q_a^2}{2a}\left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon_a}\right) + \frac{q_b^2}{2b}\left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon_b}\right) + \frac{1}{2L}\sum_{l>0}\left(q_b\mathcal{Q}_l^a + q_a\mathcal{Q}_l^b\right). \tag{20}$$

For $l=0$ and $l=1$, one can easily use either Eq. (17) or (19) to find Q_l^a and Q_l^b for use in Eq. (20) for the energy. For $l=0$ the energy is simply

$$U = \frac{q_a q_b}{\epsilon_0 L} + \frac{q_a^2}{2a}\left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon_a}\right) + \frac{q_b^2}{2b}\left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon_b}\right),$$

the interaction energy plus the two so-called self-energies. For the purpose of calculating the force, only the term that depends on L is important. Evidently, in the lowest order, the force is the same as the force for two points charges in a medium of dielectric constant ϵ_0.

For $l=1$, we let $\mathcal{Q}_{l>1}^{a,b}=0$:

$$\mathcal{Q}_1^a = \frac{(1-x_a)t_a^3}{2+x_a}\left[\frac{q_b}{\epsilon_b} + \mathcal{Q}_0^b + 2\mathcal{Q}_1^b\right],$$

$$\mathcal{Q}_1^b = \frac{(1-x_b)t_b^3}{2+x_b}\left[\frac{q_a}{\epsilon_a} + \mathcal{Q}_0^a + 2\mathcal{Q}_1^a\right],$$

whence

$$\mathcal{Q}_1^a = \frac{\epsilon_0^{-1}\alpha_1^a(q_b + q_a\alpha_1^b\beta_{11})}{1-\beta_{11}^2\alpha_1^a\alpha_1^b},$$

$$\mathcal{Q}_1^b = \frac{\epsilon_0^{-1}\alpha_1^b(q_a + q_b\alpha_1^a\beta_{11})}{1-\beta_{11}^2\alpha_1^a\alpha_1^b}.$$

Therefore, for $l=1$ the energy is

$$U = \frac{q_a q_b}{\epsilon_0 L} + \frac{q_a^2}{2a}\left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon_a}\right) + \frac{q_b^2}{2b}\left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon_b}\right) + \frac{1}{2L}\left(q_b\mathcal{Q}_1^a + q_a\mathcal{Q}_1^b\right), \tag{21}$$

with $\mathcal{Q}_1^a$ and $\mathcal{Q}_1^b$ given by the preceding two equations.

C. Effect of the dielectric constant of the spheres

According to Eq. (20), the energy of two dielectric spheres, each with a charge at the center, embedded within in infinite external dielectric medium depends in a complicated manner on the dielectric constants of the spheres. However, precise values for the dielectric constants of the objects that the spheres are to represent might not be readily available. In order to assess the impact of the dependence of the interaction on the magnitude of the dielectric response of the spheres, the energy and force of the interaction have been calculated for two identical spheres of unit radius and with unit positive charges embedded in an infinite medium of dielectric constant 80 (approximately that of water). The calculations were repeated for two values of the dielectric constant within the spheres: 4 (similar to one estimate of the dielectric constant of protein [2]) and 1 (vacuum). The difference in energy for these two internal dielectric constants [see Fig. 3(a)] is almost entirely due to the Born-like solvation terms in the energy. These terms do not depend on the separation of the two spheres. Consequently, as shown in Fig. 3(b), the force does not depend strongly on the choice of dielectric constant for the spheres.

D. Comparison with the gernalized Born approximation

Because variations of the GB model are commonly used in MD simulations of molecules in water, we wish to com-

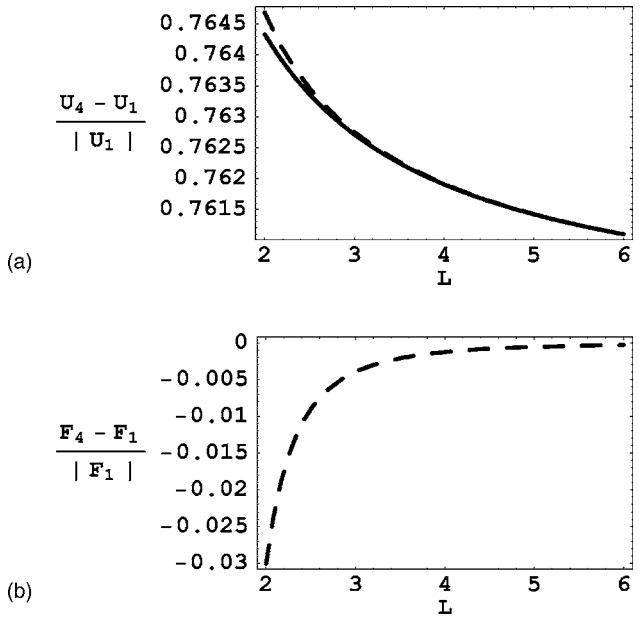

FIG. 3. Energy and force calculations for two identical spheres of unit radius with unit charges embedded in an infinite medium ($\epsilon_0=80$). (a) The dashed line is the energy for spheres with internal dielectric constant 4 minus the energy for spheres with internal dielectric constant 1 divided by the magnitude of the energy for spheres with internal dielectric constant 1. The solid line is calculated in the same manner as the dashed line except that in the numerator only the L-independent self-energy terms are kept. The solid line therefore represents the portion of the change due to the solvation self-energy terms, which do not affect the force. (b) The force for spheres with internal dielectric constant 4 minus the force for spheres with internal dielectric constant 1 divided by the magnitude of the force for spheres with internal dielectric constant 1.

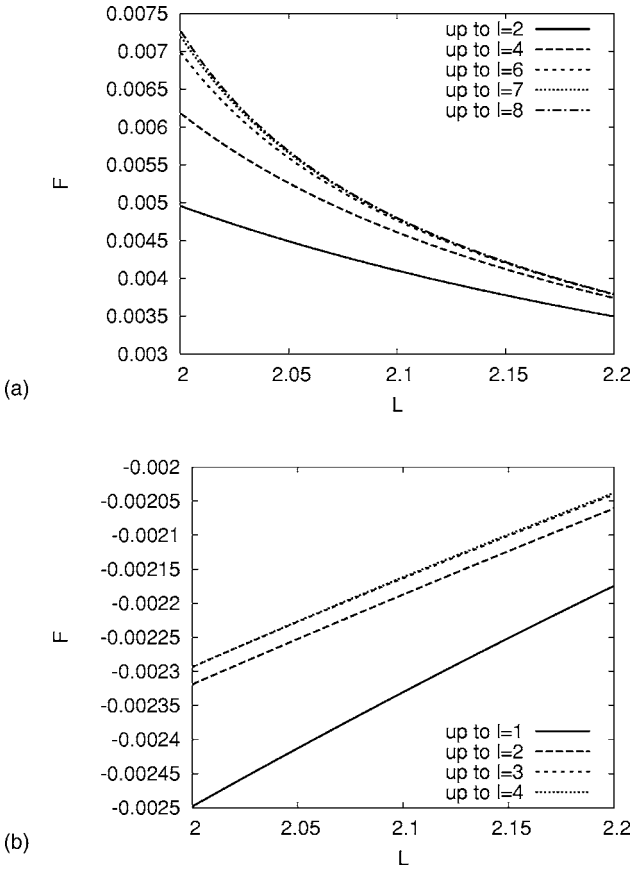

FIG. 4. The force between two spheres of unit radius, with unit interior dielectric constant, in an external dielectric with $\epsilon_0=80$. (a) The force for like charges of unit magnitude. The change between $l=7$ and $l=8$ at $L=2$ is 0.96%. The convergence is even better for larger L. (b) The force for opposite charges of unit magnitude. Excellent convergence leads to a change between $l=3$ and $l=4$ of 0.018% at $L=2$. As in the (a), the convergence is even better for larger L.

pare our result to the GB model. Since many variations of the GB model exist, we will compare to what is probably the simplest and earliest version [11]. Subsequent variations of the GB model generally take this version as a starting point for further developments in a variety of directions. In particular, for multiple spheres the GB electrostatic energy would be

$$U_{\mathrm{GB}} = -\frac{1}{2}\left(1 - \frac{1}{\epsilon_0}\right)\sum_{i,j}\frac{q_i q_j}{f_{\mathrm{GB}}} + \frac{1}{2}\sum_{i\neq j}\frac{q_i q_j}{\epsilon_0 r_{ij}}, \tag{22}$$

where the q_i $(i=1,\ldots,N)$ are the charges associated with the spheres, ϵ_0 is the dielectic constant of the solvent, and L_{ij} are the separations of the spheres. The usual form for f_{GB} is

$$f_{\mathrm{GB}}^2 = L_{ij}^2 + \alpha_i\alpha_j \exp(-L_{ij}^2/4\alpha_i\alpha_j),$$

where the α_i are the Born radii. In the version of the GB model that we compare with, the Born radii are computed according to the formula

$$\frac{1}{\alpha_i} = \frac{1}{4\pi}\int_{\mathrm{solvent}}\frac{1}{r^4}\mathrm{d}V,$$

where r is distance from the center of the ith sphere and integration takes place over the whole solvent region. The Born radii depend on the configuration of the spheres through the limits of integration, and therefore, when taking a derivative in order to obtain the forces, the Born radii must be regarded as functions of the positions of the spheres. Although this fact is sometimes neglected in the interests of computation efficiency [11], we do not neglect it in order to produce a fair comparison. Often the GB energy is given as solvation energy without the final term in Eq. (22), but since we have computed the total energy, it is necessary to include not just the solvation energy but the Coulomb energy in the comparison as well.

The comparison is made for the situation of two spheres of unit radius and unit internal dielectric constant in an infinite solvent with $\epsilon_0=80$ (water). Both opposite and like charge cases were evaluated using unit magnitude charges in both cases. Because the special case of two spheres was considered, Eqs. (19) and (20) were used to compute the energy given by the surface charge method to order $l=2$, which provides sufficient accuracy at most distances (see Fig. 4).

Figure 5 shows a graph of the force between two spheres with unit internal dielectric constant, unit radii, and an external dielectric medium with $\epsilon_0=80$. The case of equal unit

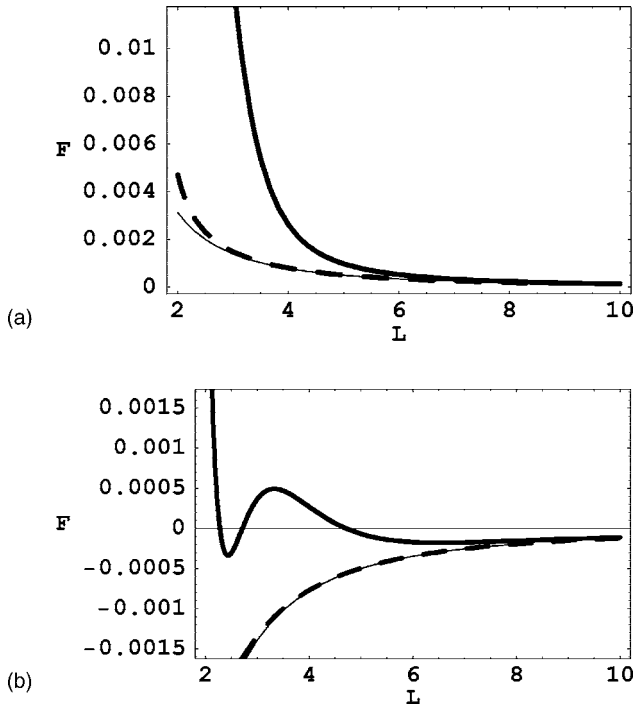

FIG. 5. Comparison of forces for unit interior dielectric constant, unit radii, and external dielectric $\epsilon_0=80$. (a) The force for like charges: the dashed line is the surface charge (exact) method ($l=2$), the heavy solid line is the GB model, and the light solid line is the result for point charges in the solvent without any spheres. (b) The force for opposite charges: the dashed line is the surface charge (exact) method ($l=2$), the heavy solid line is the GB model, and the light solid line is the result for point charges in the solvent without any spheres.

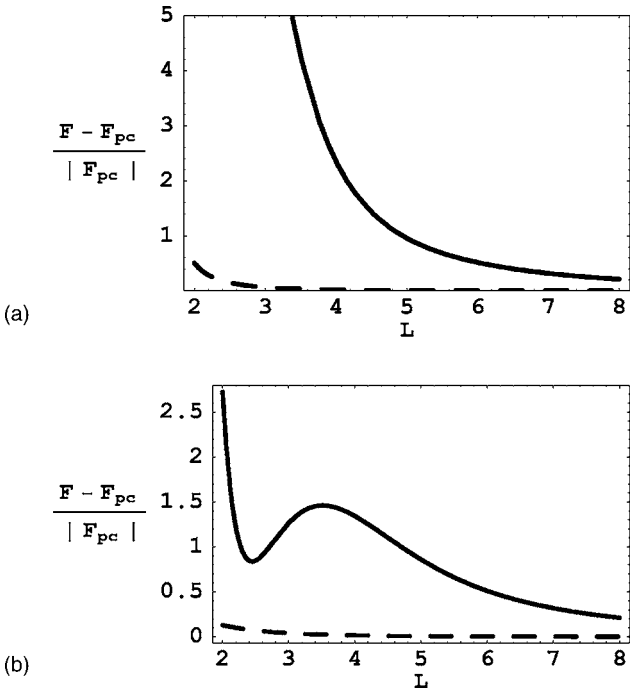

FIG. 6. Percent deviation of the force from the force of point charges in the solvent without any spheres (F_{pc}). (a) The force deviation for like charges: the dashed line is for the surface charge method ($l=2$), and the heavy solid line is for the GB model. (b) The force deviation for opposite charges: the dashed line is for the surface charge method ($l=2$), and the heavy solid line is for the GB model.

charges is in Fig. 5(a), and the case of opposite unit charges is in Fig. 5(b). Interestingly, the exact (to $l=2$) repulsion in the case of like charges is larger than the repulsion of two point charges in the solvent medium without any spheres, while the exact attraction in the case of opposite charges is smaller in magnitude than the attraction of two point charges in the solvent medium without any spheres. In the case of opposite charges, the replacement of high dielectric-constant-medium around the point charges with low-dielectric-constant material (the spheres) causes the attraction to become weaker. The same behavior occurs in the completely analytical result, to be presented in a separate publication, for symmetrically placed point charges in low-dielectric-constant material with a slab of high-dielectric-constant material in between. Possible implications of this asymmetric screening will be mentioned in the following section. The odd behavior of the GB approximation in the case of opposite charges is apparently a result of the dependence of the Born radii on the separation of the spheres. Figure 6 shows the percent difference of the exact (to $l=2$) and GB results from the force of two point charges in the solvent medium without any spheres. Some of the more elaborate variations of the GB approximation might do better than the simple one presented here, which evidently is unlikely to produce reliable results in MD simulations. Simply using the result for point charges in a solvent produces closer agreement with our exact result. However, the differences at short distances might cause potentially important differences in the local conformations.

VIII. DISCUSSION

The model presented here allows arbitrary accuracy in computing the electrostatic energy and forces in a system of charged dielectric spheres. Furthermore, the implementation of the model should be numerically efficient, since the main computational step in the computation of the energy is the inversion of a matrix of size on the order of the number of spheres. Unlike GB methods, the solute (i.e., the various spheres) need not have the same dielectric response in order for the formalism to be valid. In the special case of two isolated spheres in an infinite solvent, the exact result presented here includes an interesting asymmetric screening: in the case of like-charged spheres, the magnitude of the (repulsive) force is larger than for point charges in the same solvent, while for spheres with opposite charges, the magnitude of the (attractive) force is smaller than for point charges in the same solvent. Because this asymmetric screening appears to be a generic phenomenon not restricted to the specific geometry presented here, the asymmetric screening penalizes unfavorable contacts in a conformation of a molecule and diminishes the depth of the attractive well that other parts of the molecule might encounter when a unfavorable contact

occurs elsewhere, thereby reducing the effect of energetic traps. We conjecture that this asymmetric screening might therefore play a role in enhancing recognition of motifs in intermolecular interactions and aid in selection of the biologically functional conformation within molecules.

As alluded to earlier, the model solved here also provides electrostatic interactions in an electrorheological fluid. Because such a system offers the quickest and cleanest route to apply the model, we intend to test the model in that context first and present the details in a later publication. The restriction to a system of spheres is not a problem for a colloidal system, but can be a problem in the MD context. However, no difficulty exists in applying the surface charge method in a boundary element formulation which should scale well. In particular, given an arbitrary shape for a molecule, the surface can be divided into small patches, each with a certain surface charge. The boundary condition is applicable at each patch and will yield algebraic equations for the values of the surface charge of each patch. The presence of ions should also be accounted for and will be addressed in a future publication.

ACKNOWLEDGMENTS

This research was supported by the Intramural Research Program of the NIH, NLM. We acknowledge Morgan Varner for significant help in numerical tabulation of the matrix elements in the Appendix.

APPENDIX: CHANGE OF COORDINATES FOR FUNCTIONS ON A SPHERE

It is now necessary to find the relationship between σ^i_{lm} and $\tilde{\sigma}^j_{lm}$. We may temporarily drop the superscript i without causing confusion. We also switch briefly to Dirac notation to simplify the calculation. Namely, the ket $|\sigma\rangle$ represents the function σ, the ket $|\theta,\phi\rangle$ represents the direction given by θ and ϕ, the value of the function evaluated at the point θ,ϕ is represented by the number $\langle\theta,\phi|\sigma\rangle$, and so forth. The ket $|l,m\rangle$ is a spherical harmonic function, and $\langle\theta,\phi|l,m\rangle = Y_{lm}(\theta,\phi)$. Suppose that $\hat{n}$ is a unit vector specified by θ and ϕ. A rotation R of the vector $\hat{n}$ in three-dimensional space yields another vector $\hat{\tilde{n}}=R\hat{n}$ specified by $\tilde{\theta}$ and $\tilde{\phi}$. The corresponding rotation of a ket is represented by $\mathcal{D}(R)$: namely, $|\tilde{\theta},\tilde{\phi}\rangle=\mathcal{D}(R)|\theta,\phi\rangle$. We note that $\mathcal{D}^\dagger(R)=\mathcal{D}^{-1}(R)=\mathcal{D}(R^{-1})$.

Using the completeness of the spherical harmonics, which implies that the identity operator I may be written as $I=\Sigma_{lm}|l,m\rangle\langle l,m|$, one finds $|\sigma\rangle=\Sigma_{lm}|l,m\rangle\langle l,m|\sigma\rangle$, and therefore $\sigma(\theta,\phi)=\Sigma_{lm}Y_{lm}(\theta,\phi)\sigma_{lm}$, which is the first equality of Eq. (3). However, the identity operator may also be written as

$$I=\sum_{lm}\mathcal{D}^\dagger(R)|l,m\rangle\langle l,m|\mathcal{D}(R),$$

from which one finds

$$\sigma(\theta,\phi)=\langle\theta,\phi|\sigma\rangle=\sum_{lm}\langle\theta,\phi|\mathcal{D}^\dagger(R)|l,m\rangle\langle l,m|\mathcal{D}(R)|\sigma\rangle$$

$$=\sum_{lm}Y_{lm}(\tilde{\theta},\tilde{\phi})\tilde{\sigma}_{lm},$$

where $\tilde{\theta}$ and $\tilde{\phi}$ are the rotated coordinates. Therefore, $\tilde{\sigma}_{lm}$ becomes

$$\tilde{\sigma}_{lm}=\langle l,m|\mathcal{D}(R)|\sigma\rangle=\sum_{l'm'}\langle l,m|\mathcal{D}(R)|l',m'\rangle\sigma_{l'm'},$$

which is the desired relation between σ_{lm} and $\tilde{\sigma}_{lm}$ provided the matrix elements $\langle l,m|\mathcal{D}(R)|l',m'\rangle$. These matrix elements are known [24] and vanish unless $l=l'$. The most general rotation can be parametrized as $R(\alpha,\beta,\gamma)=R_z(\alpha)R_y(\beta)R_z(\gamma)$ where the rotations on the right-hand side are rotations about the individual axes (indicated in the subscripts) in the global coordinate system. Clearly, $R^{-1}(\alpha,\beta,\gamma)=R_z(-\gamma)R_y(-\beta)R_z(-\alpha)$. The remaining matrix elements are

$$\mathcal{D}^{(l)}_{m,m'}(R)=\langle l,m|\mathcal{D}(R)|l,m'\rangle=e^{-im\alpha-im'\gamma}\sum_{k=0}^{l+m'}(-1)^{k-m'+m}$$
$$\times\frac{\sqrt{(l+m')!\,(l-m')!\,(l+m)!\,(l-m)!}}{k!\,(l+m'-k)!\,(l-k-m)!\,(k-m'+m)!}$$
$$\times\left(\sin\frac{\beta}{2}\right)^{2k-m'+m}\left(\cos\frac{\beta}{2}\right)^{2l-2k+m'-m}. \quad \text{(A1)}$$

The sum runs over values of k for which the factorials are defined.

In order to apply Eq. (A1) to the situation at hand, the appropriate rotation R must be specified by choosing α, β, and γ. In the development above, application of R to the functions takes them from the global system to the local system. Therefore, R takes the local z axis to the global z axis. Equivalently, R^{-1} takes the global z axis to the local z axis. Inspection of Fig. 1 implies the assignments $\alpha=0$, $-\beta=\pi-\vartheta$, and $-\gamma=\varphi-\pi$. Since α vanishes, the rotation is parametrized by just two angles. This fact may be understood by noticing that the orientation of the new x and y axes is irrelevant, and therefore no rotation is necessary to bring them into a preferred position. The matrix elements $\mathcal{D}^{(l)}_{m,m'}$ for two special cases were used in Sec. VII: $\alpha=0$, $\beta=0$, $\gamma=0$ and $\alpha=0$, $\beta=\pi$, $\gamma=0$. In the first case, the factor $[\sin(\beta/2)]^{2k-m'+m}$ will vanish unless its exponent vanishes. This condition picks out a single value in the sum over k—namely, $k=(m'-m)/2$. The requirement that all the arguments in the factorial expressions be non-negative implies that $m=m'$. Consequently, $\mathcal{D}^{(l)}_{m,m'}(\alpha=0,\beta=0)=\delta_{mm'}$. Similarly, $\mathcal{D}^{(l)}_{m,m'}(\alpha=0,\beta=\pi)=(-1)^{l+2m}\delta_{m,-m'}$.

[1] Edited by D. R. Lide *CRC Handbook of Chemistry and Physics* (CRC Press, Boca Raton, 2003).
[2] B. H. Honig, W. L. Hubbell, and R. F. Flewelling, Annu. Rev. Biophys. Biophys. Chem. **15**, 163 (1986).
[3] R. MacKinnon, FEBS Lett. **555**, 62 (2003).
[4] J. Tomasi and M. Persico, Chem. Rev. (Washington, D.C.) **94**, 2027 (1994).
[5] C. J. Cramer and D. G. Truhlar, Chem. Rev. (Washington, D.C.) **99**, 2161 (1999).
[6] D. Bashford and D. A. Case, Annu. Rev. Phys. Chem. **51**, 129 (2000).
[7] E. L. Mehler, Theor. Comput. Chem. **3**, 371 (1996).
[8] S. A. Hassan, F. Guarnieri, and E. L. Mehler, J. Phys. Chem. **104**, 6478 (2000).
[9] M. K. Gilson, M. E. Davis, B. A. Luty, and J. A. McCammon, J. Phys. Chem. **97**, 3591 (1993).
[10] M. Born, Z. Phys. **1**, 45 (1920).
[11] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson, J. Am. Chem. Soc. **112**, 6127 (1990).
[12] M. Schaefer and C. Froemmel, J. Mol. Biol. **216**, 1045 (1990).
[13] M. Schaefer and M. Karplus, J. Phys. Chem. **100**, 1578 (1996).
[14] M. S. Lee, F. R. Salisbury, Jr., and C. L. Brooks III, J. Chem. Phys. **116**, 10606 (2002).
[15] A. Ghosh, C. S. Rapp, and R. A. Friesner, J. Phys. Chem. B **102**, 10983 (1998).
[16] Y.-K. Yu, Physica A **326**, 522 (2003).
[17] T. P. Doerr and Y.-K. Yu, Am. J. Phys. **72**, 190 (2004).
[18] P. B. Shaw, Phys. Rev. A **32**, 2476 (1985).
[19] R. Allen, J.-P. Hansen, and S. Melchionna, Phys. Chem. Chem. Phys. **3**, 4177 (2001).
[20] R. J. Zauhar and R. S. Morgan, J. Mol. Biol. **186**, 815 (1985).
[21] A. Yethiraj and A. van Blaaderen, Nature (London) **421**, 513 (2003).
[22] R. Friedberg and Y.-K. Yu, Phys. Rev. B **46**, 6582 (1992).
[23] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic Press, New York, 1963).
[24] J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley, Reading, MA 1994).