

# Fighting slow electrostatic Approximations.

W. Andreas Svrcek-Seiler

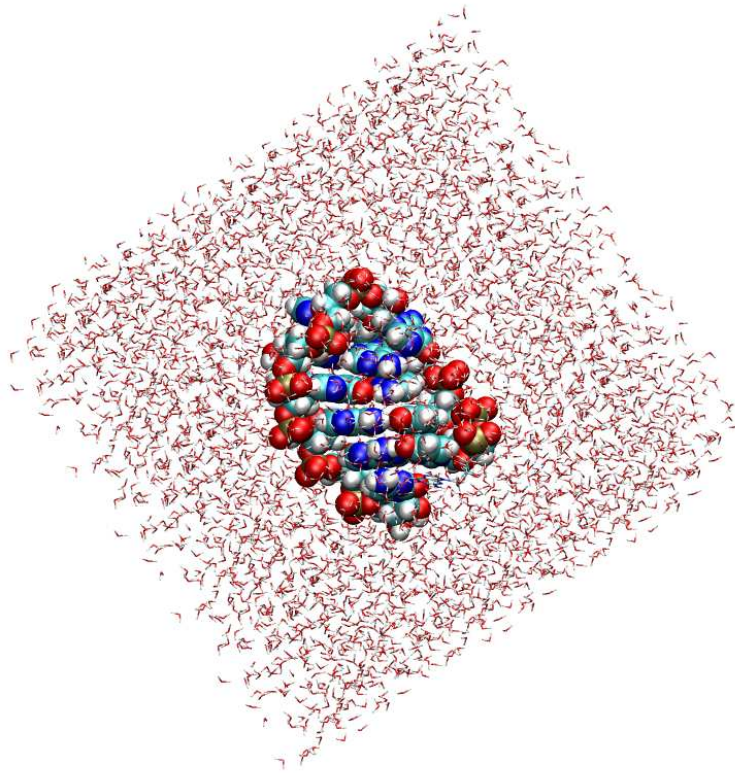
Institute for Theoretical Chemistry and Structural Biology  
University Vienna

<http://www.tbi.univie.ac.at/~svrci/>

*Bled, Feb. 2005*

Why do we care ?

*A more or less* accurate description of the interactions between molecules and their surroundings is crucial for simulating their behavior in solution. The most accurate method today is to explicitly represent surrounding water molecules and ions.



...439 RNA atoms, 10797 water atoms (hydrogen and oxygen).

## Approximations:

1. Molecules are modeled as classical systems.
2. Surrounding water is treated as dielectric continuum.
3. The effect of the dielectric continuum is again approximated as a pairwise function of the solute interatomic distances.

## Molecules as classical systems

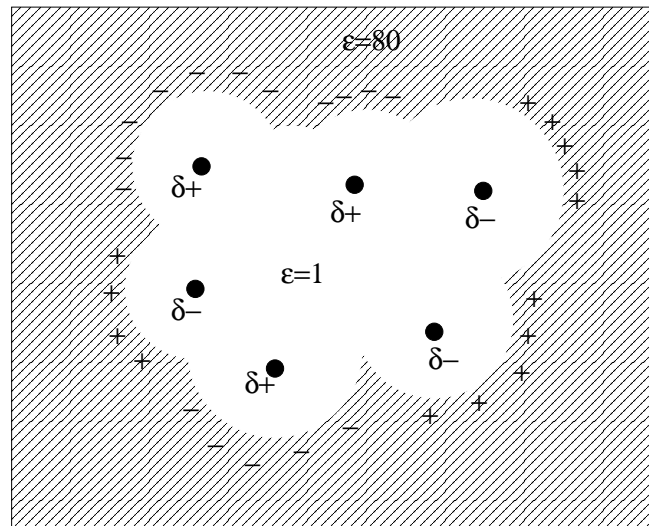
Molecular behavior is determined by a classical potential:

$$\begin{aligned} E_{\text{total}} &= \sum_{\text{bonds}} K_r (r - r_{eq})^2 \\ &+ \sum_{\text{angles}} K_\theta (\theta - \theta_{eq})^2 \\ &+ \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \\ &+ \underbrace{\sum_{i < j} \left[ \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]}_{O(N^2)\text{-consumes 99\% of CPU time}} \end{aligned}$$

...also termed **force field**, though it is a scalar (but forces can be derived from it).

## Continuum electrostatics

The charges (atoms) polarize the surrounding dielectric. The non-uniform polarization of the dielectric give rise to a potential of it's own - the **reaction field**  $\Phi_{\text{RF}}$  and this field interacts with the charges.



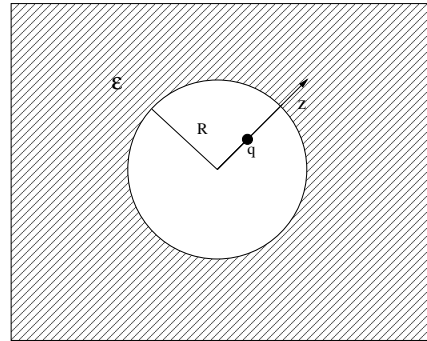
## Exact continuum electrostatics

The electrostatic potential for any arrangement of charges and dielectric can be obtained by solving the poisson equation:

$$\nabla[\varepsilon(\vec{r})\nabla\Phi(\vec{r})] = -4\pi\rho(\vec{r})$$

This is usually solved numerically (very time consuming), or analytically for some special situations.

## Kirkwood for the clueless



One begins with an off-center charge inside a spherical vacuum bubble ( $\epsilon = 1$ ), surrounded by an infinite dielectric medium. Then one can make the following ansatz for the electrostatic potential inside ( $\Phi_{int}$ ) and outside ( $\Phi_{ext}$ ):

$$\Phi_{int}(\vec{r}) = \frac{q}{|\vec{r} - d\vec{e}_z|} + \underbrace{\sum_{l=0}^{\infty} A_l r^l P_l(\cos(\theta))}_{\Phi_{RF}}$$

$$\Phi_{ext}(\vec{r}) = \sum_{l=0}^{\infty} B_l r^{-(l+1)} P_l(\cos(\theta))$$



Result:

$$\Phi_{RF}(r, \theta) = -\left(1 - \frac{1}{\epsilon}\right) \frac{q}{R} \sum_{l=0}^{\infty} \frac{(dr)^l}{R^{2l}} \frac{l+1}{l(1+1/\epsilon)+1} P_l(\cos(\theta))$$

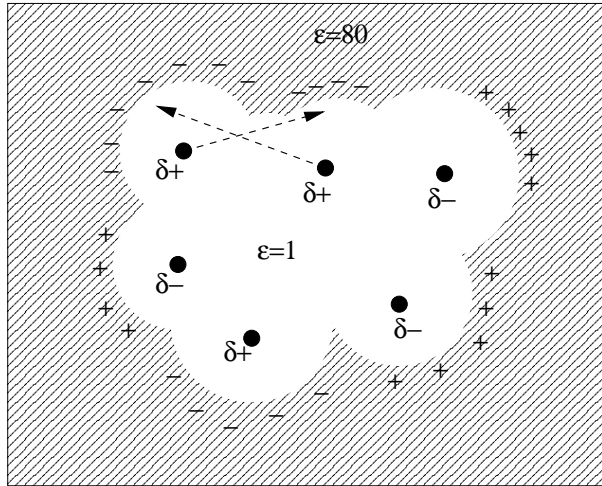
The “solvation energy” for charge  $q$  is the given by

$$E_S = \frac{q}{2} \Phi_{RF}(\vec{r} = \vec{r}_q)$$

The “solvation energy” for an arbitrary number of charges now simply follows from the superposition principle:

$$E_S(q_1, \vec{r}_1, \dots, q_N, \vec{r}_n) = -\left(1 - \frac{1}{\epsilon}\right) \frac{1}{2R} \sum_{l=0}^{\infty} \frac{l+1}{l(1+1/\epsilon)+1} \sum_{i,j}^N q_i q_j \left(\frac{r_i r_j}{R^2}\right)^l P_l(\cos(\theta_{ij}))$$

## Baseline:



- The solvation energy contains  $q_i^2$  terms, which correspond to the hydrophilicity of charged atoms and  $q_i q_j$  terms, leading to a roughly attractive force component between “like charge” atoms.
- For a single charge  $q$  at the center of the sphere, one obtains the famous Born formula:

$$E_S = -\left(1 - \frac{1}{\epsilon}\right) \frac{q^2}{2R}$$

- For a single charge in off-center position, one obtains the closes expression

$$E_S(d) = -\frac{q^2}{2R} \frac{R^2}{R^2 - d^2}$$

...in the limit  $\epsilon \rightarrow \infty$ .

## From Born to generalized Born (*GB*)

The Born formula can be obtained by a simple volume integral:

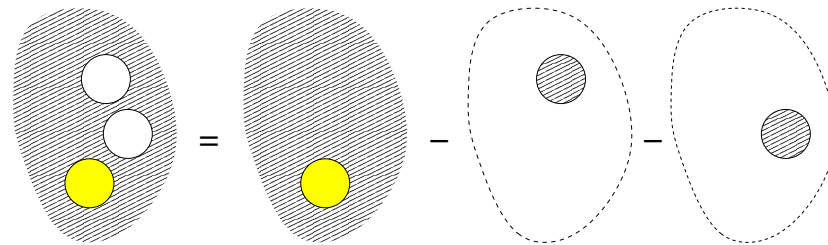
$$E_S = -\frac{1}{8\pi} \int_{\text{outside}} (\epsilon - 1) \vec{E} \cdot \vec{E}_0 dV$$

To some extent, this result holds also if the charge is not located at the center of the cavity. Within the *GB* scheme, this is inverted to

$$\frac{1}{a_i} = \frac{1}{4\pi} \int_{\text{outside}} \frac{dV}{|\vec{r} - \vec{r}_i|^4}$$

The volume integral is usually performed over atomic volumes as in:

$$\int_{\text{outside}} \frac{dV}{r^4} = \int_{r>r_i} \frac{dV}{r^4} - \sum_{j \neq i}^N \int_{V_j \setminus V_i} \frac{dV}{r^4}$$



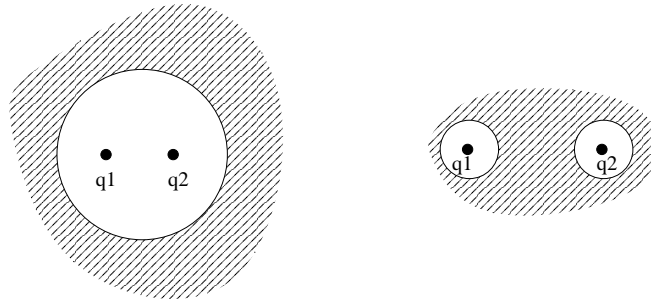
The per-atom contributions are known as functions of atomic radii and interatomic distance.

The total energy of the system now becomes:

$$\begin{aligned}
 E_{\text{total}} &= \sum_{\text{bonds}} K_r (r - r_{eq})^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_{eq})^2 \\
 &+ \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \\
 &+ \sum_{i < j} \left[ \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{R_{ij}} - \left(1 - \frac{1}{\epsilon}\right) \frac{q_i q_j}{f^{\text{GB}}(R_{ij}, a_i, a_j)} \right] - \left(1 - \frac{1}{\epsilon}\right) \sum_i \frac{q_i^2}{2a_i}
 \end{aligned}$$

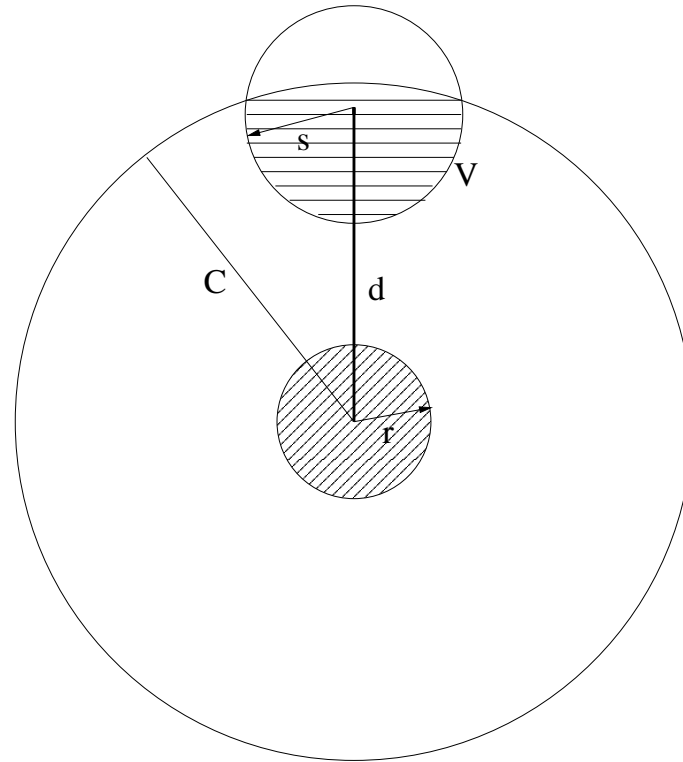
with  $f^{\text{GB}}$  commonly chosen as

$$f^{\text{GB}} = [r_{ij}^2 + a_i a_j \exp(-r_{ij}^2 / 4a_i a_j)]^{\frac{1}{2}}$$



## Locating bottlenecks and speeding things up **(I)**

The analytical expressions for the volume integral to calculate the *effective radii*  $a_i$  are and involve *all* atom pairs  $i, j$ . Solution: perform the volume integral over a restricted volume:



## Speeding things up **(II)**

The most time-consuming step is to evaluate

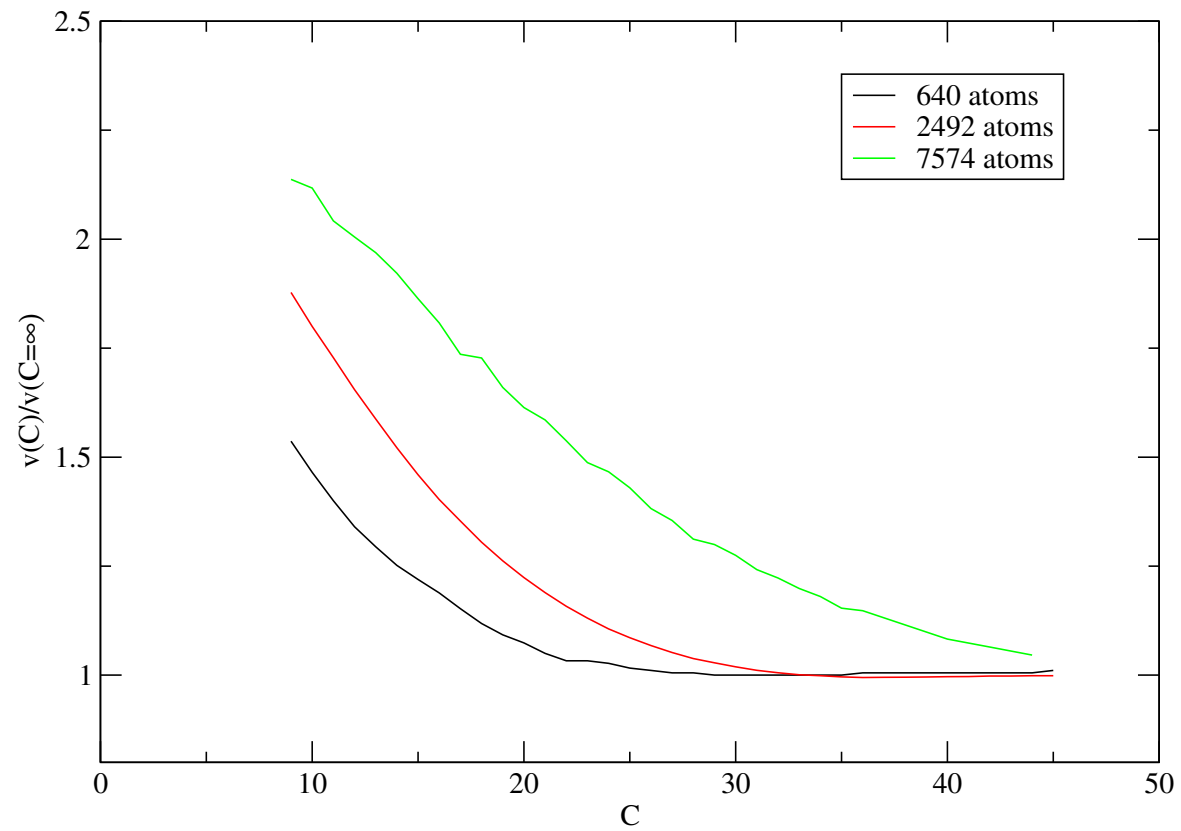
$$X = \frac{1}{2} \left( \frac{s}{d^2 - s^2} + \frac{1}{2d} \log \frac{d - s}{d + s} \right)$$

with  $d \gg s$  in most cases. If  $d \gg s$ ,

$$X = \frac{s^3}{d^4} \left( \frac{1}{3} + \frac{2s^2}{5d^2} + \frac{3s^4}{7d^4} + O\left(\frac{s^6}{d^6}\right) \right)$$

...This series around “ $\infty$ ”, truncated after 4 terms is sufficiently accurate and a lot faster on common CPUs.

# Result



...Thorough tests showed no deterioration of the trajectories!



## Conclusions:

- We significantly improved the performance of simulations based on the GB model.
- The reliability of the results does not deteriorate due to our trick (thanks partly to the dubious quality of the model in itself).
- Researchers can produce up to twice as much data output within the same time.

Appendix: It could have been worse:

After solving the poisson equation one gets for the potential 'outside' a spherical cavity:

$$\Phi_{out}(r, \theta) = \frac{q}{\epsilon r} \sum_{l=0}^{\infty} \left(\frac{d}{r}\right)^l \frac{2l+1}{l(1+\frac{1}{\epsilon})+1} P_l(\cos(\theta))$$

Though there's a much more straightforward way, we now calculate the 'solvation energy' for  $q$  through a volume integral over the 'outside' region as:

$$E_s = -\frac{1}{8\pi} \int_{out} (\epsilon - 1) \vec{E} \vec{E}_0 dV \quad (1)$$

$\vec{E}_0$  is the gradient of the  $\Phi$  without dielectric,  $\vec{E}$  the gradient of  $\Phi_{out}$  from above.

Remembering the gradient for polar coordinates, one gets for  $\vec{E}\vec{E}_0$ :

$$\vec{E}\vec{E}_0 = \frac{q^2}{\epsilon} \left( \frac{\partial\Phi}{\partial r} \frac{\partial\Phi_0}{\partial r} + \frac{1}{r^2} \frac{\partial\Phi}{\partial\theta} \frac{\partial\Phi_0}{\partial\theta} \right)$$

This expands to ( $\frac{q^2}{\epsilon}$  omitted):

$$\begin{aligned} \vec{E}\vec{E}_0 &\propto \left( \sum_{l=0}^{\infty} \frac{d^l}{r^{l+2}} (l+1) P_l(\cos(\theta)) \right) \cdot \\ &\quad \cdot \left( \sum_{l=0}^{\infty} \frac{d^l}{r^{l+2}} \frac{(2l+1)(l+1)}{l(1+\frac{1}{\epsilon})+1} P_l(\cos(\theta)) \right) + \\ &\quad + \left( \sum_{l=0}^{\infty} \frac{d^l}{r^{l+2}} P'_l(\cos(\theta)) \sin(\theta) \right) \cdot \\ &\quad \cdot \left( \sum_{l=0}^{\infty} \frac{d^l}{r^{l+2}} \frac{2l+1}{l(1+\frac{1}{\epsilon})+1} P'_l(\cos(\theta)) \sin(\theta) \right) \end{aligned}$$

One sees that this expression involves the sum of two products of infinite series, which seems quite cumbersome. However, the  $P_l$  s are orthogonal and I found a nice orthogonality relation for the  $P'_l$  s too:

$$\int_{-1}^1 P'_m(x)P'_n(x) \cdot (1 - x^2)dx = \frac{2n(n + 1)}{2n + 1}\delta_{m,n}$$

Using the orthogonality relations one can put things back together and the solvation energy  $E_s$  becomes:

$$E_s = -\frac{1}{8\pi}\left(1 - \frac{1}{\epsilon}\right)4\pi q^2 \int_R^\infty \sum_{l=0}^\infty \frac{d^{2l}}{r^{2l+4}} \frac{(2l + 1)(l + 1)}{l\left(1 + \frac{1}{\epsilon}\right) + 1} r^2 dr$$

The rest is straightforward and one arrives again at the 'one charge Kirkwood' formula:

$$E_s = -\frac{1}{2}\left(1 - \frac{1}{\epsilon}\right)q^2 \sum_{l=0}^\infty \frac{d^{2l}}{R^{2l+1}} \frac{l + 1}{l\left(1 + \frac{1}{\epsilon}\right) + 1}$$

...This was done to solve a physics puzzle (but didn't lead anywhere) :-)

Thanks

...For your attention and patience .  
Further thanks go to Alexey Onufriev.