

1. Introduction

1.1 Electrostatics

- Electrostatic potential
- differential equation
- Coulomb sum
- Bjerrum length

We assume N charges $q_i e$ at positions r_i . This corresponds to a charge density

$$\rho e = \sum_{i=1}^N q_i e \delta(r - r_i). \quad (1.1)$$

If the charges move slowly, then we can ignore magnetic forces, and the force on each particle is

$$F_i = q_i e E(r_i) \quad (1.2)$$

and the total electrostatic energy

$$U_E = \int \rho(r) e E(r) dr, \quad (1.3)$$

where E is the electrostatic field with the properties

$$\nabla \cdot E = \rho e / \epsilon \quad (1.4)$$

$$\nabla \times E = 0. \quad (1.5)$$

The latter implies that E is the gradient of an up to a constant unique scalar electrostatic potential Φ such that $E = -\nabla\Phi$. Inserting into (1.4) gives

$$-\nabla^2 \Phi = \rho e / \epsilon, \quad (1.6)$$

which is a defining differential equation for Φ . For a single charge, it has the solution:

$$-\nabla^2 \frac{q_i e}{4\pi\epsilon} \frac{1}{r - r_i} = q_i e \delta(r - r_i) / \epsilon \quad (1.7)$$

It is convenient to define the *Bjerrum length*

$$l_B = \frac{e^2}{4\pi\epsilon k_B T}, \quad (1.8)$$

which is the distance at which two unit charges feel an electrostatic interaction equal to $1k_B T$; this means for example that by spontaneous thermal fluctuations, two oppositely charged unit charges can be separated by a distance up to l_B . The Bjerrum length in water at room temperature is 0.7nm due to its large relative permittivity of $\epsilon/\epsilon_0 \approx 80$.

Then the electrostatic forces and potential are

$$F_i = k_B T l_B \sum_{j \neq i} \frac{q_i q_j}{|r_i - r_j|^3} (r_i - r_j) \quad \text{and} \quad U_E = \frac{1}{2} k_B T l_B \sum_{i=1}^N \sum_{j \neq i} \frac{q_i q_j}{|r_i - r_j|}, \quad (1.9)$$

where the (infinite) self interactions of point charges are omitted. The differential equation for Φ is

$$-\nabla^2 \Phi = 4\pi \frac{l_B k_B T}{e} \rho. \quad (1.10)$$

In the following, the Bjerrum length and $k_B T$ will be omitted.

1.2 Boundary conditions and conditional convergence

- periodic boundary conditions
- different summation orders
- convergence factors
- dipole term

To avoid finite size effects, computer simulations are typically performed with periodic boundary conditions. This means, we consider an infinite set of charges $q_i e$ at positions $r_i + n$, where

$$n = (kl_x, ll_y, ml_z) \quad \forall k, l, m \in \mathbb{Z} \quad (1.11)$$

and l_x, l_y, l_z are the periodic cell sizes in the three spatial dimensions. In the following, we use the short notation $n \in l\mathbb{Z}^3$, and $V = l_x \cdot l_y \cdot l_z$ denotes the volume of the periodic cell. We assume *charge neutrality*, that is

$$\sum_{i=1}^N q_i = 0. \quad (1.12)$$

We consider the electrostatic energy of the N primary charges with all others, that is

$$U_E = \frac{1}{2} \sum_{i=1}^N \sum_{n \in l\mathbb{Z}^3} \sum_{j=1}^N{}' \frac{q_i q_j}{r_i - r_j - n}, \quad (1.13)$$

where $'$ denotes that for $n = 0$, the $j = i$ -term is omitted. While being the obvious extension of Eqn. (1.9) to a periodic system, this is in fact not a well defined equation. The reason is that the sum is not absolutely convergent; the value of U_E depends on the order of summation over k, l, m . Since we assume that our N particles are representative for the bulk of a large, macroscopic sample that is approximated by the periodic sum, a reasonable choice is a summation in radially ascending shells:

$$U_E = \frac{1}{2} \sum_{i=1}^N \sum_{S=0}^{\infty} \sum_{n^2=S} \sum_{j=1}^N{}' \frac{q_i q_j}{r_i - r_j - n}. \quad (1.14)$$

Different summation orders are possible, which result in different values for U_E . In addition, U_E can be defined via different routes. For example, one can solve the differential Eqn. (1.10) under periodic boundary conditions, then

$$U_E = \frac{1}{2} \sum_{i=1}^N q_i \Phi(r_i). \quad (1.15)$$

This definition of the electrostatic potential is called *intrinsic*. Finally, one can use *convergence factors* to define U_E :

$$U_E := \lim_{\beta \rightarrow 0} \frac{1}{2} \sum_{i=1}^N \sum_{S=0}^{\infty} f(S, \beta) \sum_{n^2=S} \sum_j{}' \frac{q_i q_j}{r_i - r_j - n}, \quad (1.16)$$

where the convergence factor $f(S, \beta)$ has the following properties:

1. $f(S, \beta)$ is a continuous function for $\beta \geq 0$
2. $f(S, 0) = 1$
3. $f(S+1, \beta) \leq f(S, \beta)$ for all S, β
4. $0 \leq f(S, \beta) \leq 1$ for all S, β .

That is, *formally* we recover the original series for $\beta = 0$. However, if at fixed $\beta > 0$, $f(S, \beta)$ decays fast enough (for example exponentially), then

$$\sum_{S=0}^{\infty} f(S, \beta) \sum_{n^2=S} \sum_{j=1}^N \frac{q_i q_j}{r_i - r_j - n}, \quad (1.17)$$

is absolutely convergent and well defined, so that also the limit (1.16) is well defined.

The convergence factor approach is the most general one, since it allows to obtain the same results as the other approaches; we will use this approach to prove some of the methods for evaluating the Coulomb sum. For example, the convergence factor $e^{-\beta S^2}$ corresponds to summation in ascending spherical shells. The convergence factor $e^{-\beta|r+S|^2}$ leads to the intrinsic potential.

However, the difference between summation orders is not as big as expected. Smith has shown that for example

$$U_{\text{spherical}} = U_{\text{intrinsic}} + \frac{2\pi}{3V} M^2 \quad (1.18)$$

$$U_{\text{slabwise}} = U_{\text{intrinsic}} + \frac{2\pi}{V} M_z^2, \quad (1.19)$$

where $U_{\text{spherical}}$ denotes the summation result for summation in ascending shells, U_{slabwise} diskwise summation in two dimensions, and after that symmetric summation in the remaining dimension, and $U_{\text{intrinsic}}$ the intrinsic solution.

Note that the *dipole term* $\frac{2\pi}{3V} M^2$ has in fact a physical interpretation, if written as

$$\frac{2\pi}{(2\epsilon'_r + 1)V} M^2, \quad (1.20)$$

where ϵ'_r is the dielectric permittivity of a medium that is supposed to surround the crystal (remember that we interpret the crystal as growing in ascending shells). For vacuum, $\epsilon'_r = 1$, so that we recover the term as in Eqn. (1.18). Another important value are metallic boundary conditions, for which the term vanishes. In case the system contains free ions, this is the only reasonable boundary condition, since otherwise this term imposes a harmonic potential that drives the particles back to an imaginary central cell.

Note also that for this term the unfolded positions r_i matter. For calculating the energy (1.13), it seems that the position r_i only matter up to multiples of l_x , l_y and l_z . After all, we are investigating a large, periodic crystal, and shifting r_i for example by l_x in the x -direction, the internal structure of the crystal will not change. However, the dipole term will change; moreover, it will be discontinuous, even if particle coordinates are modified only by multiples of l_* . Therefore, it is important to use unfolded, that is continuous coordinates for its calculation.

Instead of three dimensional periodicity, one can also consider two- or onedimensionally periodic systems; this is used for studying inter-/surfaces or wires and nanotubes. One can also generalize the periodicity to non-orthorombic simulation boxes. Another problem arises when for example calculating solvation free energies. In this case, a single charge is added to the system, which consequently becomes non-neutral. As one can easily, the energy of a non-neutral system is necessarily infinite. However, one can assume a homogeneous neutralizing background, which being homogeneous does not generate forces on the particles. Therefore, the dynamics of the system is unbiased by the background. But to calculate the energy, this background is necessary, since otherwise the result depends on the applied method, and even on non-physical tuning parameters.

Literature

- S. W. de Leeuw, J. W. Perram and E. R. Smith, Proc. R. Soc. Lond. A 373, 1980.
E. R. Smith, Proc. R. Soc. Lond. A 375, 1981. J. D. Jackson, *Classical Electrodynamics*, Wiley, 1999.

2 Ewald method

The Ewald method is *the* classical method for calculating electrostatic interactions in computer simulations with periodic boundary conditions. It dates back to 1921 and was originally invented by Ewald to calculate crystal energies, especially the Madelung constant (by hand, not by computer, of course!). By carefully tuning the method parameters, the method nevertheless has a computational scaling of $N^{3/2}$ and is therefore significantly faster than a plain all-with-all N^2 -force calculation. Nowadays, there are several other methods scaling like $N \log N$ or even N , but all are much less elegant than the Ewald methods, and have rather large overheads, so that for small systems, Ewald's method is still a good choice. Moreover, unlike its mesh-extensions, the Ewald method converges exponentially, that is to double the amount of valid digits in the result, one just needs to double the computational effort.

The main trick of the Ewald method is to split the two main problems of the Coulomb potential sum - the singularities at the positions of the charges, and the slow decay of the potential, which makes it conditionally convergent. To this aim, the delta distributions that represent the charge densities of the charges are replaced by Gaussian charge clouds of width α^{-1} :

$$q_i \delta(r - r_i) = q_i [\delta(r - r_i) - \rho_{\text{Gauss}}(r - r_i)] + q_i \rho_{\text{Gauss}}(r - r_i) \quad (2.1)$$

where the Gaussian cloud is defined as

$$\rho_{\text{Gauss}}(r) = \left(\frac{\alpha}{\sqrt{\pi}} \right)^3 e^{-\alpha^2 r^2}. \quad (2.2)$$

To make use of this splitting, we calculate the electrostatic energy via the electrostatic potential, that is as

$$U_E = \frac{1}{2} \sum_{j=1}^N q_j \Phi(r_j), \quad (2.3)$$

where

$$\Phi(r) = \sum_{j=1}^N \sum'_{n \in \mathbb{Z}^3} \frac{q_j}{r - r_j - n}. \quad (2.4)$$

The ' on the sum denotes that the (infinite) terms with $r = r_j$ are omitted.

Fourier space sum

Although the sum over the Gaussians still leads to an only conditionally convergent sum, we can however solve the differential equation for the electrostatic potential, since unlike the original charge distribution, the Gaussians are smooth, and (1.10) can be solved for a periodic and smooth solution. Such a function can be conveniently represented as Fourier series:

$$f(r) = \frac{1}{l^3} \sum_{k \in \frac{2\pi}{l} \mathbb{Z}} \hat{f}(k) e^{ikr}, \quad (2.5)$$

where

$$\hat{f}(k) = \int_V dr f(r) e^{-ikr} \quad (2.6)$$

is the Fourier transform of f . Note that for simplicity we assume a cubic periodic cell V of size $l \times l \times l$. Since for any smooth periodic function

$$\nabla_r f(r) = \frac{1}{l^3} \sum_{k \in \frac{2\pi}{l}\mathbb{Z}} \hat{f}(k) \nabla_r e^{ikr} = \frac{1}{l^3} \sum_{k \in \frac{2\pi}{l}\mathbb{Z}} ik \hat{f}(k) e^{ikr}, \quad (2.7)$$

Poisson's equation for the Fourier series of the electrostatic potential Φ transforms into

$$k^2 \hat{\Phi}(k) = 4\pi \hat{\rho}(k). \quad (2.8)$$

We now consider the charge density given by the sum of all Gaussians, which has the Fourier series

$$\hat{\rho}_G(k) = \int_V dr e^{-ikr} \sum_{j=1}^N q_j \sum_{n \in l\mathbb{Z}^3} \left(\frac{\alpha}{\sqrt{\pi}} \right)^3 e^{-\alpha^2(r-r_j-n)^2} = \sum_{j=1}^N q_j e^{-ikr_j} e^{-k^2/4\alpha^2} \quad (2.9)$$

Note that

$$\hat{\rho}_G(0) = \int_V dr \sum_{j=1}^N q_j \sum_{n \in l\mathbb{Z}^3} \left(\frac{\alpha}{\sqrt{\pi}} \right)^3 e^{-\alpha^2(r-r_j-n)^2} = \sum_{j=1}^N q_j = 0 \quad (2.10)$$

due to charge neutrality; therefore, Eqn. (2.8) trivially holds, but does not allow to determine $\Phi(0)$. However, normalization requires that this constant term is 0. By insertion of Eqn. (2.9) into (2.8), we obtain the electrostatic potential due to the Gaussians Φ_G , and from this the interaction of the charges with these Gaussians:

$$U_F = \frac{1}{2} \sum_{j=1}^N q_j \Phi_G(r_j) = \frac{1}{2} \sum_{j=1}^N q_j \sum_{k \neq 0} e^{-ikr_j} \frac{4\pi}{k^2} \hat{\rho}_G(k) = \frac{1}{2V} \sum_{k \neq 0} \left| \sum_{i=1}^N q_i e^{ikr_j} \right|^2 \frac{4\pi}{k^2} e^{-k^2/4\alpha^2}. \quad (2.11)$$

Real space sum

What remains is to calculate the contribution due to the interaction between the charges and the real, discrete charges and the charge-inverted Gaussians. The potential of the inverted Gaussians rapidly screens the potential of the point-like charges, so that the sum of both over the lattice is absolutely and quickly convergent; therefore, we can perform it efficiently in real space.

For this, we need is the potential also due to a single Gaussian. Again, we use Poisson's equation, but this time in real space:

$$-\frac{1}{r} \frac{\partial^2 r \Phi_{\text{Gauss}}(r)}{\partial r^2} = 4\pi \rho_{\text{Gauss}}. \quad (2.12)$$

Integration on both sides gives

$$-\frac{\partial r \Phi_{\text{Gauss}}(r)}{\partial r} = - \int_r^\infty dr 4\pi r \rho_{\text{Gauss}} = - \left(\frac{\alpha}{\sqrt{\pi}} \right)^3 \int_r^\infty dr 4\pi r e^{-\alpha^2 r^2} = -2 \frac{\alpha}{\sqrt{\pi}} e^{-\alpha^2 r^2}. \quad (2.13)$$

And a second integration gives

$$\Phi_{\text{Gauss}}(r) = \frac{2\alpha}{\sqrt{\pi} r} \int_0^r dr e^{-\alpha^2 r^2} = \frac{\text{erf}(\alpha r)}{r}, \quad (2.14)$$

where erf denotes the error function.

We now compute the real space energy

$$U_R = \frac{1}{2} \sum_{i \neq j} q_j q_i \left[\frac{1}{|r_i - r_j|} - \Phi_{\text{Gauss}}(|r_i - r_j|) \right] + \frac{1}{2} \sum_i q_i^2 \Phi_{\text{Gauss}}(0) + \quad (2.15)$$

$$\frac{1}{2} \sum_{i,j} q_j q_i \sum_{n \neq 0} \left[\frac{1}{|r_i - r_j - n|} - \Phi_{\text{Gauss}}(|r_i - r_j - n|) \right] \quad (2.16)$$

where the first line are the contributions due to the primary image $n = 0$, in which we have to omit the singular self-energy of the point charges. We need therefore to evaluate two terms:

$$\Phi_{\text{Gauss}}(0) = \lim_{r \rightarrow 0} \Phi_{\text{Gauss}}(r) = \frac{\alpha}{\sqrt{\pi}} \quad (2.17)$$

and

$$\frac{1}{r} - \Phi_{\text{Gauss}}(r) = \frac{\text{erfc}(\alpha r)}{r} \quad (2.18)$$

The complementary error function erfc decays exponentially fast. By choosing α sufficiently large, one can obtain U_R easily by a summation only over nearest images of each pair of charges.

The total (intrinsic) Coulomb energy is then approximately given by

$$U_E = U_{\text{intrinsic}} \approx \frac{1}{2V} \sum_{0 < |k|^2 < K^2} \left| \sum_{i=1}^N q_i e^{ikr_j} \right|^2 \frac{4\pi}{k^2} e^{-k^2/4\alpha^2} \quad (2.19)$$

$$+ \frac{1}{2} \sum_{\substack{i \neq j \\ |r_i - r_j|_{\text{m.i.}} < r_c}} q_i q_j \frac{\text{erfc}(\alpha |r_i - r_j|_{\text{m.i.}})}{|r_i - r_j|_{\text{m.i.}}} \quad (2.20)$$

$$- \frac{1}{2} \frac{\alpha}{\sqrt{\pi}} \sum_i q_i^2, \quad (2.21)$$

where we have introduced the computationally necessary truncations in k -space by cutoff K and in real space by cutoff r_c . The force on particle i can be easily obtained by taking the derivative with respect to r_i , since all sums are absolutely and fast converging.

2.1 Error estimates

What are now optimal values for K , R and α ? Or given values, what is the numerical error we make? From Eqn. (2.19), it is easy to see that the real space cutoff should be inversely proportional to α , while the k -space cutoff should be proportional. This allows us to estimate the computational effort of the Ewald summation: Using a cell list-like approach and assuming homogeneously distributed particles, the computation time for the evaluation of the real space sum is Nr_c^3 . For the k -space, the computation time is NK^3 . Inserting the expected scalings with α , we get a total computation time $N\alpha^{-3} + N\alpha^3$, which is minimal for $\alpha = N^{1/6}$. Therefore, the optimal scaling of the Ewald sum is $N^{3/2}$.

However, this does not yet allow to reliably choose the parameters; for this, we need the error estimates by Kolafa and Perram. For the error estimate, we assume that the charges are homogeneously randomly distributed. Our goal is to calculate the root mean square (RMS) error

$$\Delta F = \sqrt{\langle (F_{\text{exact}} - F_{\text{Ewald}})^2 \rangle} = \sqrt{\langle (F_{\text{exact}}^{\text{real}} - F_{\text{Ewald}}^{\text{real}})^2 \rangle + \langle (F_{\text{exact}}^{\text{Fourier}} - F_{\text{Ewald}}^{\text{Fourier}})^2 \rangle}, \quad (2.22)$$

assuming that the real and Fourier space errors are uncorrelated.

By replacing the sum by an integral, the real space error can be estimated as

$$\Delta F^{\text{real}} \approx \frac{\sum q_i^2}{\sqrt{N}} \frac{2}{\sqrt{r_c V}} e^{-\alpha^2 r_c^2}, \quad (2.23)$$

and the Fourier error as

$$\Delta F^{\text{Fourier}} \approx \frac{\sum q_i^2}{\sqrt{N}} \frac{2\alpha}{\sqrt{\pi K V}} e^{-K^2/4\alpha^2}. \quad (2.24)$$

This error estimates do not only allow to estimate the error at given K , r_c and α , but can also be used to tune the Ewald sum optimally for given error goal ΔF . To this aim, one loops over Fourier-cutoffs K and determines the minimal α numerically from Eqn. (2.24), such that the RMS error is

below $\sqrt{2}\Delta F$. Then, Eqn. (2.23) allows to determine the minimal cutoff r_c , so that the real space error is also below $\sqrt{2}\Delta F$. Now, one just needs to try through a couple of Fourier space cutoffs to see which combination of K and r_c is optimal for the given implementation and hardware.

2.2 Neutralization

As mentioned before, it is possible to consider systems with a net charge by assuming a homogeneous background that fills the entire simulation and neutralizes the system. This background does not exert a force due to symmetry, but it contributes to the energy; the contribution in the case of the Ewald summation is

$$U_N = -\frac{\pi}{2\alpha^2 V} \left(\sum_{i=1}^N q_i \right)^2, \quad (2.25)$$

which can for example be easily calculated by integration the implicit pair interaction

$$U_{ij} = \frac{1}{V} \sum_{0 < |k|^2 < K^2} q_i q_j e^{ik(r_i - r_j)} \frac{4\pi}{k^2} e^{-k^2/4\alpha^2} + q_i q_j \sum_{n \in \mathbb{Z}^3} \frac{\operatorname{erfc}(\alpha|r_i - r_j - n|)}{|r_i - r_j - n|}, \quad (2.26)$$

of which only the second sum gives the above term.

This contribution depends on α , therefore it must not be neglected when calculating the energy of nonneutral systems.

Literature

- Frenkel and Smit, *Understanding Molecular Simulation*, Academic Press, 2002.
 S. W. de Leeuw, J. W. Perram and E. R. Smith, Proc. R. Soc. Lond. A 373, 1980.
 Kolafa and Perram, Mol. Sim. 9, 1992.

3 mesh-accelerated Ewald methods (P^3M , SPME)

- parallelization

4 Tree codes and fast multipole methods

5 MMM, MMM2D, MMM1D

6 Electrostatic Layer Correction (ELC)

7 Maggs' method (MEMD)