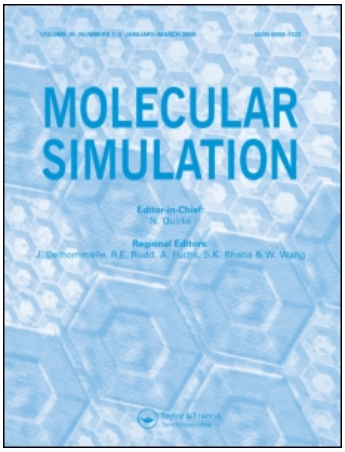

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Calculation of the Coulomb Interactions in Condensed Matter Simulation

Z. A. Rycerz[ab]
[a] Central Research Laboratory, Hitachi Ltd., Tokyo, Japan [b] Department of Chemistry, The University of Western Ontario, London, Ontario, Canada

To cite this Article Rycerz, Z. A.(1992) 'Calculation of the Coulomb Interactions in Condensed Matter Simulation', Molecular Simulation, 9: 5, 327 — 349
To link to this Article: DOI: 10.1080/08927029208049125
URL: http://dx.doi.org/10.1080/08927029208049125

Molecular Simulation, 1992, Vol. 9(5), pp. 327–349
Reprints available directly from the publisher
Photocopying permitted by license only

CALCULATION OF THE COULOMB INTERACTIONS IN CONDENSED MATTER SIMULATION

Z.A. RYCERZ[1]

Central Research Laboratory, Hitachi Ltd., Tokyo 185, Japan

(Received June 1992, accepted July 1992)

A method order of N for molecular dynamics (MD) simulation of multi-component ionic systems is described. The accuracy of various MD results and efficiency of simulation are discussed in terms of the range of the r-space and q-space Ewald sums. It has been shown that for large enough systems (order of 10^3 particles) the q-space sum can be neglected without worsening the quality of MD results. Its ommission additionally accelerates the $O(N)$ method described by one to two orders of magnitude and make it possible to simulate on a mainframe the systems containing order of 10^6 particles.

KEY WORDS: Ewald sums, ionic systems, molecular dynamics.

1 INTRODUCTION

Molecular Dynamics simulation plays at present an important role in studies of fundamental processes in many areas of research. Among others it is employed extensively in the study of chemical reactions, catalysis, various physical processes in condensed matter, ion beams in heavy-ion accelerators, in plasma physics, quantum chromodynamics, formation of star clusters in astrophysics and in the structural investigation of complex biological (DNA) and chemical systems. The critical point for many of those applications is the size of the system that can be simulated in a reasonable time on presently available computers. Some applications require the simulation of the systems containing as many as 10^5–10^6 particles. Examples in the field of condensed matter are the study of the dynamics of a low concentration of impurities (or defects) in solids, the calculation of physical quantities which involve correlations over a large distance r, the study of hydrodynamic instability at the microscopic level, the study of the process of melting, the simulation of complex materials of technological importance like glassy films, GaAs semiconductors and high-temperature oxide superconductors, and others. A classical $O(N^2)$ MD algorithm cannot be used to investigate large systems due to the enormous amount of CPU time required. For example, the simulation of a 10^6 particle system would require several years of CPU time on the most powerful presently available supercomputers. Consequently several existing problems cannot be tackled without significant improvement (by orders of magnitude) in the efficiency of existing MD codes.

In general in an MD system one may distinguish two kind of interactions: short-range interactions and long-range (coulombic type) interactions. Short-range interactions can be effectively limited to some closest neighbours contained in the

[1]Present address: Department of Chemistry, The University of Western Ontario, London, Ontario, Canada N6A 5B7

cut-off sphere of radius R_c while long-range interactions have to be calculated for all particles in the system. In fact, due to the small size of the systems usually simulated, there is a need to include the interactions with *image* particles outside the basic MD box. The Ewald method [1] is the most suitable and accurate method for the calculation of the long-range contribution to the forces in an MD simulation. In recent years considerable effort has been devoted to improving the calculation of both, the short-range and long-range interactions. Examples of new MD codes for the fast calculation of short-range interactions are those described in papers [2] to [8]. The most advanced among them enables one to simulate MD systems with short-range forces at a speed of 50 000 p/s (particles per CPU second) on a supercomputer. On the other hand the speed of coulombic part of simulation is lower by orders of magnitude; depending on the range of the q-space summation and the system's size, this speed is between 50 and 1000 p/s on the state-of-the-art supercomputers (see for example [9], [10]). Consequently any significant improvement in the efficiency of the simulation of ionic systems can be achieved only by a speed-up of the coulombic part of simulation. In recent years several papers on the calculation of coulombic interactions have been published (e.g. [9] to [14]). Up to this time the most efficient algorithm for the simulation of systems involving both, short-range and long-range interactions, seems to be that published by Perram, Petersen and De Leeuw [9] (referred to here as the PPL algorithm). The authors use the Jacobi theta function transformation to derive rapidly computable forms for the energy and forces. As in the Ewald method, that transformation also gives two sums: one in the r-space and the other in the q-space. By an appropriate choice of some parameters the over-all CPU time of the PPL algorithm grows proportionally to $N^{3/2}$, and therefore for large N their algorithm becomes much more efficient than a classical $O(N^2)$ algorithm (e.g. [15]).

In this paper recent improvements regarding the efficiency of MD simulations are discussed. The method presented can be used for the simulation of multi-component ionic systems involving both short-range and long-range interactions and its over-all CPU time is strictly proportional to the total number of particles in the system. In addition, both parts are calculated very efficiently which results in the great speed of the algorithm. The short-range interactions are calculated by a method based on the SP (scalar pyramid) $O(N)$ algorithm [6, 8]. Further (and substantial) speed-up of the $O(N)$ coulombic part of the simulation can be achieved by omission of the q-space term in the Ewald sum [10]. A justification for this omission is discussed in details in the following sections.

2 CALCULATION OF LONG-RANGE COULOMB FORCES

Because of the slow convergence of the electrostatic potential its calculation cannot be limited to the MD box only. Apart from the interactions between the charges within the MD box it is necessary to include interactions with the charges contained in image cells. The total coulombic energy of such a system (including the interactions with image particles) is given by [16],

$$\Phi^c = \frac{1}{2} \sum_{i,j}^{N\prime} \sum_{\mathbf{n}} \frac{Z_i Z_j}{r_{i,j\mathbf{n}}} \tag{1}$$

where $r_{i,j\mathbf{n}} = |\mathbf{r}_{j\mathbf{n}} - \mathbf{r}_i|$ is a distance between a particle i from the origin cell (i.e.

MD box; $\mathbf{n} = (0, 0, 0)$) and a particle j from $\mathbf{n} = (n_x, n_y, n_z)$ image cell. The prime in the first double sum denotes the omission of terms with $i = j$ when $\mathbf{n} = (0, 0, 0)$. Ewald [1] showed that the slowly convergent series (1) can be replaced by two rapidly convergent series, one in real space and the other in reciprocal space (q-space), so that

$$\Phi^c = \Phi^r + \Phi^q + \Phi^O \tag{2}$$

with the r-space term,

$$\Phi^r = \frac{1}{2}\sum_{i,j}^{N}{}' \sum_{\mathbf{n}} Z_i Z_j \frac{erfc(\alpha r_{i,j\mathbf{n}})}{r_{i,j\mathbf{n}}} \tag{3}$$

with the q-space term,

$$\Phi^q = \frac{2\pi}{L^3}\sum_{i,j}^{N} Z_i Z_j \sum_{\mathbf{q}}{}' A(q)\cos(\mathbf{q}\cdot\mathbf{r}_{ij}) \tag{4}$$

and constant term,

$$\Phi^O = -\frac{\alpha}{\sqrt{\pi}}\sum_{i=1}^{N} Z_i^2 \tag{5}$$

The factor $A(q)$ is

$$A(q) = -\frac{1}{q^2}\exp\left(-\frac{q^2}{4\alpha^2}\right) \tag{6}$$

and $erfc(x)$ denotes the complementary error function,

$$erfc(x) = 1 - \frac{2}{\sqrt{\pi}}\int_0^x e^{-u^2}\,du \tag{7}$$

α is the Ewald convergence parameter, $\mathbf{q}$ is a reciprocal space vector (i.e. for a cubic cell of side L, $\mathbf{q} = (2\pi/L)$ (k, l, m), with k, l, m integers), $q = |\mathbf{q}|$ and the prime in equation (4) denotes $\mathbf{q} \neq (0, 0, 0)$. The constant term (5) does not play a practical role in MD simulations since it gives zero contribution to the forces; however its contribution to the energy is significant and it has to be taken into account in order to get the correct value of the energy. Differentiation of eqs. (3) and (4) gives the r-space and q-space electrostatic forces.

3 DESCRIPTION OF THE METHOD

The r-space forces are calculated in the same part of the program as that which calculates the short-range forces and both employ the SP method [6, 8]. The CPU time for this part of the simulation is proportional to the average number of nearest neighbours (NN) NN_{av} which is roughly equal to $\frac{4}{3}\pi R_c^3\rho$ (R_c is the radius of the cut-off sphere). Its value per time step and per particle does not depend on the size of the system and equals,

$$t_{sr} = (t_{int} + t_{ass}/\mathrm{NTUPDA}) \times NN_{av} \approx (t_{int} + t_{ass}/\mathrm{NTUPDA}) \times \tfrac{4}{3}\pi R_c^3\rho \tag{8}$$

where: t_{int} and t_{ass} are the CPU times (which are constant for a given computer)

needed for the calculation of an interaction (distance, force, potential . . .) and for the assignment of nearest neighbours, respectively. Both times refer to a *single interaction* and both are for short-range forces, including the r-space Coulomb term. NTUPDA is the frequency of the NN list updating.

Although equation (4) involves a double summation over i and j (and thus an $0(N^2)$ operation), the CPU time for the q-space force calculation can also be changed from $0(N^2)$ into $0(N)$ dependency by the use of the trigonometrical identity:

$$\cos(\mathbf{q}\cdot\mathbf{r}_{ij}) = \cos(\mathbf{q}\cdot\mathbf{r}_j)\cdot cos(\mathbf{q}\cdot\mathbf{r}_i) + \sin(\mathbf{q}\cdot\mathbf{r}_j)\cdot\sin(\mathbf{q}\cdot\mathbf{r}_i) \tag{9}$$

To avoid an expensive calculation of trigonometric functions (sin, cos) the program stores these functions for all particle positions ($\mathbf{r}_i$ or $\mathbf{r}_j$) and all $\mathbf{q}$ values and uses them to calculate the components of the q-space force and energy. A further speed-up of this part of the simulation has been achieved by the use of additional symmetry that exists between $\mathbf{q}$ vector components and by the use of the *even* and *odd* properties of trigonometric functions. As the result the number of computer operations involved has been greatly reduced.

The range of the q-space summation is determined by an integer parameter k_O. In the case of a cubic system the calculations are performed for $k^3_{max}-1$ ($l_{max}\cdot m_{max}\cdot n_{max}-1$ for a non-cubic system) image sub-cells, plus the origin one ($k=0$). Each $\mathbf{q}$ component changes from $-k_O$ to $+k_O$ (so $k_{max}=2\times k_O+1$). The value of k_O depends on a kind of substance and the size of system. Therefore the speed of the q-space part calculation depends strongly on the range of the q-space summation (value of k_{max}) but it changes linearly with N. The CPU time of the q-space (Fourier) force calculation per time step and per particle can be written as,

$$t_{fur} = t^q_{fur}\times l_{max}\cdot m_{max}\cdot n_{max} \tag{10}$$

or

$$t_{fur} = t^q_{fur}\times k^3_{max} \quad \text{for cubic symmetry} \tag{11}$$

where t^q_{fur} denotes the CPU time for the calculation of the q-space force acting between a single i–j pair and is constant for a given computer.

Consequently the CPU times for both the short-range and long-range parts grows proportionally to N. The total CPU time per time step can be simply expressed as:

$$t_{MD} = t_{tot}\times N = (t_{sr}+t_{fur})\times N \tag{12}$$

with t_{sr} and t_{fur} given by eqs. (8) and (10) respectively.

3.1 The choice of the convergence parameter α and the efficiency of simulation

The parameter α controls the distribution of the Φ^c energy (and the Coulomb forces) between the Φ^r *and* Φ^q terms. It is convenient to consider α in dimensionless units namely,

$$\alpha_L = \alpha\cdot L \tag{13}$$

where L is the side length of MD box. In a certain range of α_L the Φ^c energy is independent of the value of α_L. This region is determined by the lower α^l_L

and upper α_L^u limits and is called the energy plateau region. To get correct results from the simulation value of α_L ought to be set between these limits. Its recommended value may be estimated as,

$$\alpha_L \approx \tfrac{1}{2}(\alpha_L^l + \alpha_L^u) \tag{14}$$

Sangster and Dixon (16) showed that for a small system ($N = 216$) the value of $\alpha = 5.6$ gives an accuracy in the coulombic force calculation which is of the order of 10^{-3} and this value of α_L has been commonly adopted in the simulation of ionic systems (15). Smith (17) expressed α in terms of the short-range force cut-off radius and suggested that $\alpha \cdot R_c$ should be ~ 3.5, which corresponds to $\alpha_L \sim 7/R_{CD}$, giving for $R_{CD} = 1$ a value of α_L somewhat bigger than the value of 5.6 given in [16] (R_{CD} is a dimensionless cut-off radius equal to $R_{CD} = R_c/H$, $H = L/2$). In fact the proper value of the α_L parameter depends strongly on the size of the system N and the choice of cut-off radius. and it should be chosen with some regard for accuracy and the CPU time consumed. This problem has been investigated recently by Rycerz and Jacobs [10] and in the present paper it is discussed in great detail in the light of the accuracy of various MD results.

It has been found through intensive MD similations (performed for *AgI*, *NaCl*, Bi_2O_3 and $SrCl_2$ ionic systems with N ranging from 216 to 10^6 particles; see sec. 4) that the lower limit of α_L is mostly determined by the cut-off radius R_c while the upper limit of α_L^u is determined by the values of N and k_{max}. Appropriate relations are as follows,

$$\alpha_L^u = N^{1/3} + f(k_{max}) \tag{15}$$

$$\alpha_L^l = \frac{\alpha_1}{R_{CD}} \tag{16}$$

with $f(k_{max})$ term roughly estimated as,

$$f(k_{max}) \approx (k_{max}/k_1)^2 \tag{17}$$

(we did not attempt to estimate this dependence more accurately since it is not so important in the aim of the paper). A typical value of α_1, a parameter which defines the smallest useful value of α_L, is 4.0 ± 0.5 and it depends to some extent on the substance under consideration. A roughly estimated value of k_1 is 4.0 ± 1.0. The values of the α_1 and k_1 parameters for the ionic systems investigated are as follows,

System	α-*AgI*	*NaCl* (molten)	δ-Bi_2O_3	$SrCl_2$
α_1	4.0 ± 0.3	4.0 ± 0.5	4.3 ± 0.5	4.1 ± 0.2
k_1	4.0 ± 0.4	3.0 ± 0.4	4.5 ± 0.3	4.2 ± 0.4

(18)

In general one can formulate the following two conclusions regarding the choice of an optimal α_L parameter (Figure 1):

(a) the plateau region expands towards smaller values of α_L when the value of R_{CD} is increased
(b) the plateau region expands towards bigger values of α_L when the size of the system N is increased

These observations make it possible to manipulate the α_L, R_{CD} and k_{max} parameters in order to increase the speed of a simulation. From the efficiency

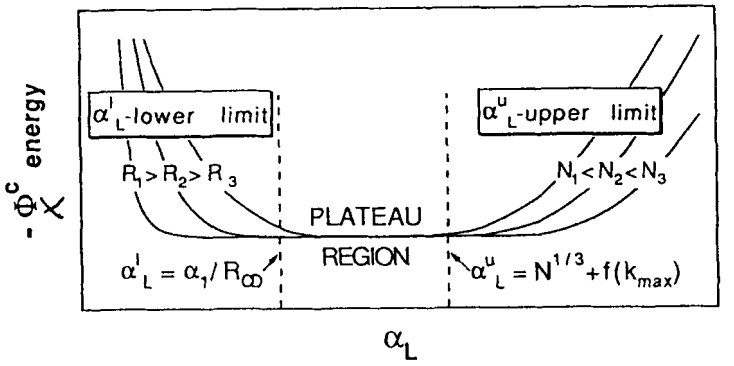

Figure 1 The range of the energy plateau region depends mainly on the choice of cut-off radius R_{CD} and on the size of an MD system N.

point of view, the cut-off radius should be as small as possible (equation (8)) and this is especially important for large MD systems. The use of a small R_{CD} requires a relatively large α_L (equation (16)) and for a system as big as $N = 10^6$ the value of α_L ought to be of the order of 100 and thus much bigger than the value of 5.6 commonly applied in the simulation of ionic systems.

The over-all speed of the algorithm is dominated by the calculation of short-range and long-range forces. This part of the simulation usually contributes more than 95% of the total CPU time. Therefore, it is important to analyze in more detail the relation between these two contributions. The values of t_{int}, t_{ass} (equation (8)) and t^q_{fur} (equation (11)) are similar on a serial computer (e.g. on the Hitac M-880 mainframe they are: 0.6, 0.8, 1.2μs, respectively). In most cases the choice of R_c such that NN_{av} is of the order of 50 to 100 is sufficient for MD purposes. Assuming those values and that the value of NTUPDA lies between 2 and 20, the ratio of t_{fur}/t_{sr} can be roughly expressed as,

$$\eta = \frac{t_{fur}}{t_{sr}} \approx \frac{k^3_{max}}{NN_{av}} \approx \frac{(2k_O+1)^3}{50} \tag{19}$$

One can see from equation (19) than even for a small value of $k_O = 2$ the majority of CPU time is consumed for the calculation of q-space forces (ca. 70%) while short-range forces requires much less CPU time ($\eta = 2.5$). For bigger k_O the ratio of η grows rapidly attaining a value close to 200 for $k_O = 10$. It is apparent then, that the omission of the q-space force calculation would result in an additional great acceleration of the MD simulation. One of the main goals of this work was to determine whether the q-space forces can be neglected in the simulation of coulombic systems. It should be noticed that the q-space term has no physical reality but was introduced into the Coulomb summation to improve its convergence only. In fact, the particles interact only in *real space* via coulombic potentials. Therefore, it seems to be possible that for large enough systems this term might be neglected in practical applications. It has been found for all the systems examined that this term can indeed, be excluded from the simulation. For large enough systems and carefully chosen α_L and R_{CD} parameters the contributions from the q-space terms to the total energy and forces can be reduced to very low levels so that their influence on the MD results becomes negligible.

Table 1 The conditions of simulations for the four ionic systems examined.

System	Range of N	T_O [K]	Phase	Used potential [ref]	ρ [(kg/m^3]	a_O [Å]	R_c range [Å]
α-AgI	256–500 000	513	Solid	[19]	5951.10	5.079	7.5–12.0
$SrCl_2$	324–12 000	1100	Solid	[20]	2821.09	3.600	7.2–12.0
δ-Bi_2O_3	270–10 000	1029	Solid	[21]	8605.14	5.644	7.5–13.7
$NaCl$	216–1 000 000	1250	Liquid	[22]	1555.13	3.148	6.1–12.0

Symbols: T_O – requested temperature of simulation, ρ – density, a_O – lattice parameter.

4 MD RESULTS FOR SELECTED IONIC SYSTEMS

The results presented have been obtained by a program based on the method described [18]. The calculation of coulombic interactions in the program is switchable, so that it can simulate both coulombic and non-coulombic systems containing from 1 to 6 species (kind of particles). For a simple one component non-coulombic system with two body potential it attains on the Hitac M-880 single processor (64 bits) the speed of up to 40 000 p/s, while for two component ionic systems its speed ranges from 5000 to 25 000 p/s, depending on the substance and conditions of a simulation.

Four kind of ionic systems have been simulated at constant volume through a wide range of N. We choose the systems that exhibit strong coulombic interactions: three superionic crystals (with highly disordered structure) and liquid $NaCl$ (the Coulomb forces in crystal $NaCl$ are much weaker due to its regular structure). The basic conditions of these simulations are collected in Table 1.

The value of the lattice parameter for $NaCl$ system given in Table 1 has been expanded by 6% to encourage melting of the constant volume system [10]. The results calculated included the following physical quantities:

(a) thermodynamic quantities:
- temperature,
- short-range potential energy,
- r-space part of Coulomb energy,
- q-space part of Coulomb energy,
- total Coulomb energy,
- total potential energy,
- kinetic energy,
- total energy,
- virial,
- internal pressure,
- Madelung constant,

(b) mean square displacement (*msd*) of atoms and diffusion coefficients,
(c) P_α – function (moments ratio for Van Hove correlation function),
(d) histogram of individual atom displacements, (e) xyz projections of the positions of atoms,
(f) $g(r)$ – partial radial distribution functions,
(g) fourth moments of the dynamic structure factor.

The simulations were carried out for 100 to 4000 time steps depending on the size of the system. The results presented are averages over the working runs length of 50 to 1000 steps. For smaller systems the simulation had been started from the restart files obtained after 5000 to 10000 steps of equilibration runs. For bigger systems the simulations were started from smaller systems in order to shorten the equilibriation process. Selected results are presented in the following sections.

4.1 Results for NaCl

In Figure 2 is shown the coulombic energy Φ vs α_L for two sizes of *NaCl* system ($N = 216$ and $N = 8000$) and for different values of R_c and k_{max}. The cut-off radius ranges from 7.5 Å to 12.0 Å and these values correspond to 56 and 232 NN in the R_c sphere, respectively. The value of k_{max} was changed from $k_{max} = 0$ (which is referred to later in the paper as *q-space coulombic interactions switched-off*) to $k_{max} = 21$ (i.e. 9260 image cells). The results show that the range of the Φ^c plateau strongly depends on all three parameters: N, R_c and k_{max}. Figure 3 shows the Φ^q term for three selected curves from Figure 2. As one can see this term is much smaller for larger systems ($N = 8000$) than for a smaller one ($N = 216$). For the small system the Φ^q terms for different k_{max} (curves 1 and 2 in Figure 3) seem to be similar; however, for bigger values of α_L their differences become very large (cf. curves 3 and 4 in Figure 2). It has been observed that the optimal value of α_L is close to $N^{1/3}$. For this reason Figure 4 shows in detail the behavior of the Φ^c energy around the $\alpha_L = N^{1/3}$ point for $N = 216$. For such a small system an energy plateau exists, indeed, only for large values of k_{max}. For smaller k_{max} it makes practically no difference whether the q-space term is switched off or on (cf. curves 2 and 3 in Figure 2). It has been found that if the calculated energy of the system departs from the plateau energy by approx. $\pm 1\%$ or more, then the differences in the MD results become significant. On the other hand if the Φ^c energy of the system is inside a $\pm 0.5\%$ (or smaller) range these differences are usually negligible. It should be taken into account that the points in Figure 4 (as in other figures) are averages over many configurations (each with slightly different value of Φ^c). If, therefore, the value of α_L is chosen so that it gives the precisely required value of Φ^c but the slope of the curve in Figure 4 is large, then the variation of Φ^c is also bigger which affects the results. When the q-space summation is switched off this slope is determined by the value of R_c and the size of the system. Figure 5 clearly shows that for $k_{max} = 0$ the use of a bigger R_c value results in a significant increase of the plateau region. This effect grows gradually for bigger systems. Even for the small value $R_c = 7.5$ Å this growth is remarkabe (Figure 6). In more detail the data from Figure 6 around the $\alpha_L = N^{1/3}$ point are shown in Figure 7. As can be seen the slope of Φ^c against $\alpha_L / N^{1/3}$ is practically constant, which means that α_L changes with the system's size as $N^{1/3}$. This dependence is shown in Figure 8. The closed circles in Figure 8 correspond to the value of energy plateau, while horizontal bars to the $\pm 0.5\%$ region (cf. Figure 7). Therefore two dashed lines may be considered as limits for the choice of an optimal value of α_L. The results of Figures 7 and 8 are not surprising. They confirm that the value of the parameter α (in physical units) is a constant for a given substance and that is what one would expect: $\alpha \equiv (\alpha_L / N^{1/3}] \times \rho^{1/3}$. This is true for a small value of R_c (for $R_c = 7.5$ Å the slope of curve 1 in Figure 4 is practically the same as that for curve 1 in Figure 5; see also Figure 7). However for bigger R_c the situation is quite different. There is a remarkable difference in the improvement of plateau region caused by the increase

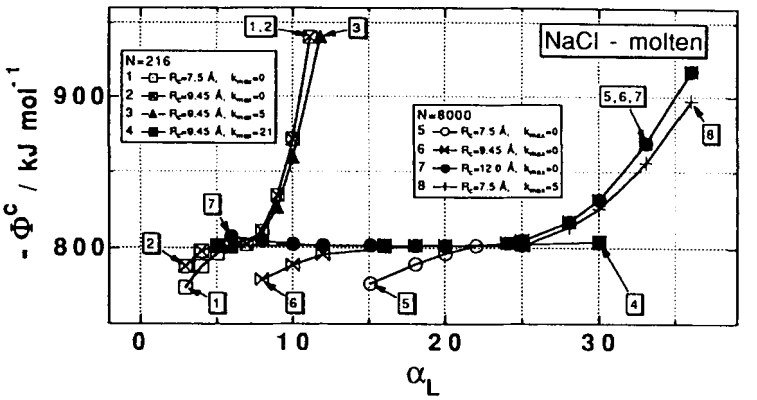

Figure 2 Coulombic energy of *NaCl* systems for different values of cut-off radius R_c and different range of q-space summation k_{max}.

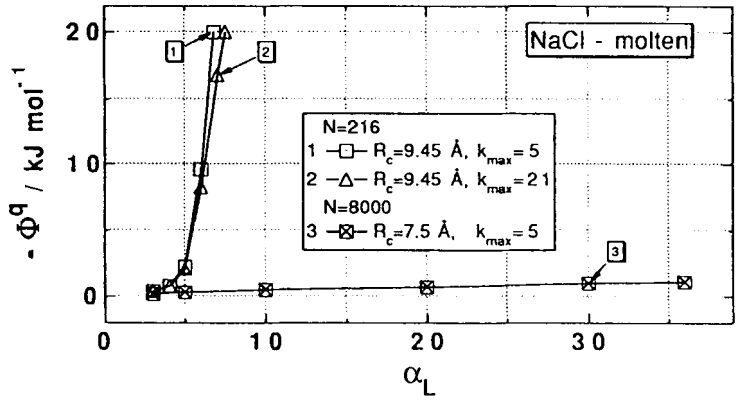

Figure 3 Energy q-space term of *NaCl* systems for different values of R_c and k_{max}.

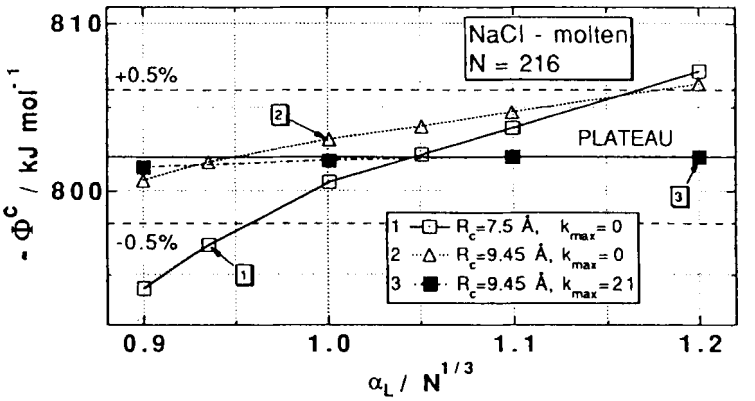

Figure 4 Coulombic energy of $N = 216$ *NaCl* system in the range of ±0.5% around the plateau level.

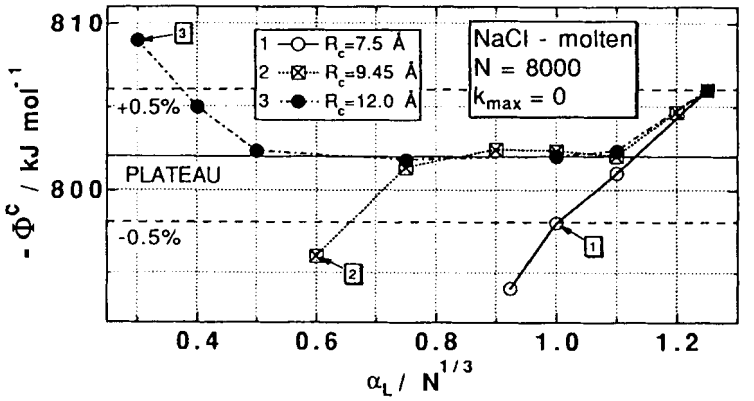

Figure 5 Coulombic energy of $N = 8000$ *NaCl* system in the range of ±0.5% around the plateau level.

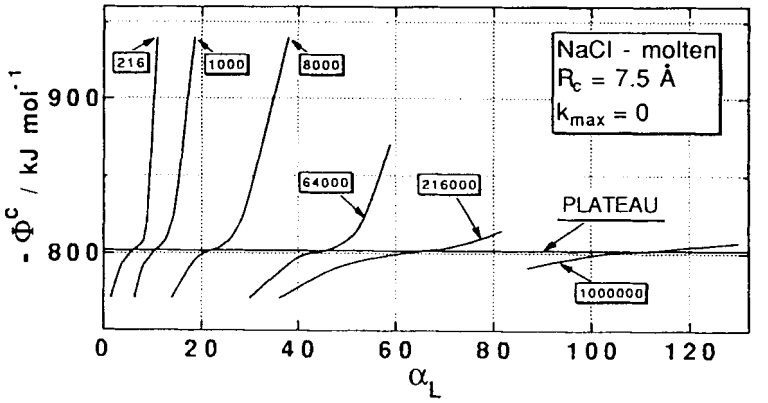

Figure 6 Coulombic energy of *NaCl* systems of different sizes (q-space term switched off).

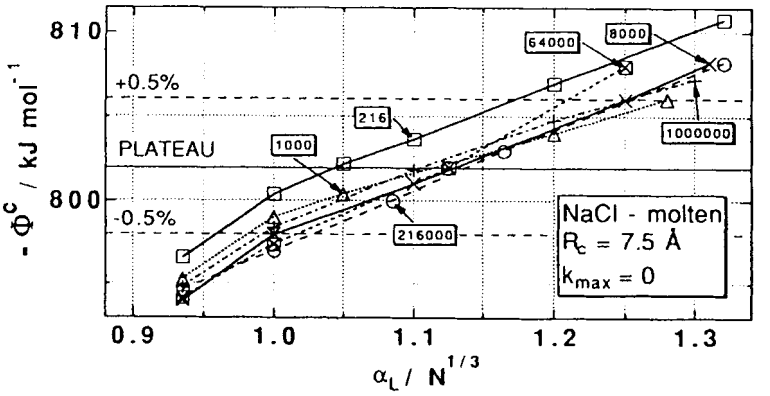

Figure 7 Coulombic energy of *NaCl* systems of different sizes in the range of ±0.5% around the plateau level (q-space term switched off).

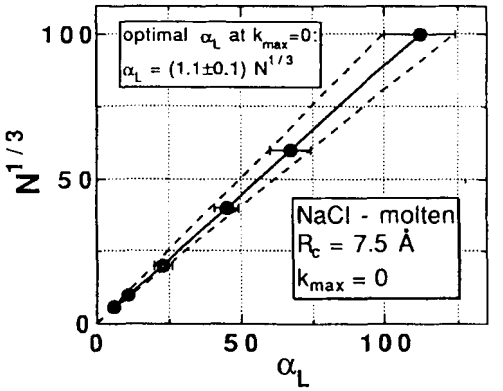

Figure 8 Dependency of the convergence parameter α_l on the system's size N. The closed circles correspond to the value of energy plateau while horizontal bars to the ±0.5% region. For more details see text.

of the cut-off radius for systems of different sizes. The use of a bigger R_c for a larger system results in the appearance of an easily distinguished flat region of energy (curve 2 in Figure 5) while similar changes for smaller system are less remarkable (curve 2 in Figure 4). This seems to be caused by two factors. The first is a simple statistical one: the averages are calculated over better equilibriated configurations (for a bigger system). The second one is related to self-correlations introduced through the periodic boundary conditions (this effect is especially important when the coulombic interactions are present). It is worthwhile noticing that this improvement is even more remarkable when the q-space term is included. The calculation of the q-space term – for a given step of simulation – consists in the repetitive summation over the same configuration of atoms but for different values of $\boldsymbol{q}$ vector. Therefore the q-space summation is performed over the configuration which is more strongly correlated for a smaller system than for a bigger one.

The conclusion of this part of the work can be formulated as follows. For a large enough system and properly chosen values of R_c and α_L the omission of the q-space term does not affect in practice the energy of the system. To verify the validity of this statement, for quantities other than the energy, several quantities (as listed at the beginning of the section) have been examined. Some of them are less sensitive on the changes in N, R_c, k_{max} and α_L values (as for example $g(r)$), while the others are much more sensitive. The most sensitive among those listed are diffusion coefficient, pressure and virial, and they have therefore been selected for presentation.

In Figures 9 and 10 are shown diffusion coefficients (calculated from *msd*) for *Na* and *Cl* ions. The internal pressure of the system is shown in Figure 11 and the virial in Figure 12. All these results were obtained with the q-space term switched-off ($k_{max} = 0$) and R_c equal to 7.5 Å. Such a small value of R_c has been chosen to emphasize sensitivity of the results to the size of the system and the value of α_L (for bigger R_c this sensitivity becomes gradually smaller). For each size of system several simulations were performed with different values of α_L. In the case of diffusion coefficient (Figures 9 and 10) the results for $R_c = 12.0$ Å of $N = 8000$ system have been presented as well for comparison with those for $R_c = 7.5$ Å. On each figure the closed circles represent the values at the energy plateau while the

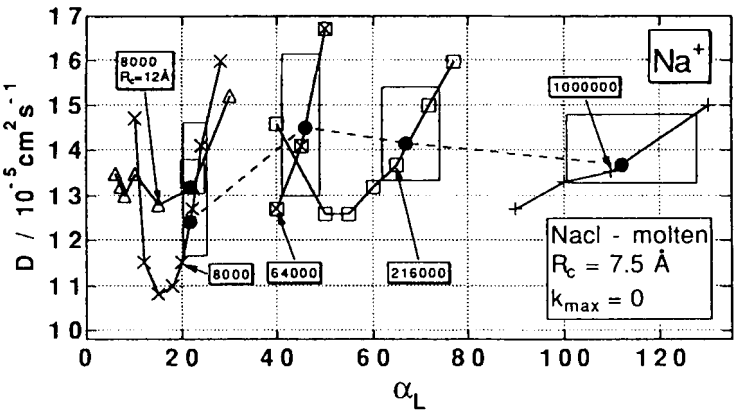

Figure 9 Diffusion coefficient of *Na* ions for *NaCl* systems of different sizes (q-space term switched off). For comparison the curve with $R_c = 12$ Å for $N = 8000$ system is given. The closed circles represent the values at the energy plateau while the rectangles around them show the area of ±0.5% range of the energy plateau. For more details see text.

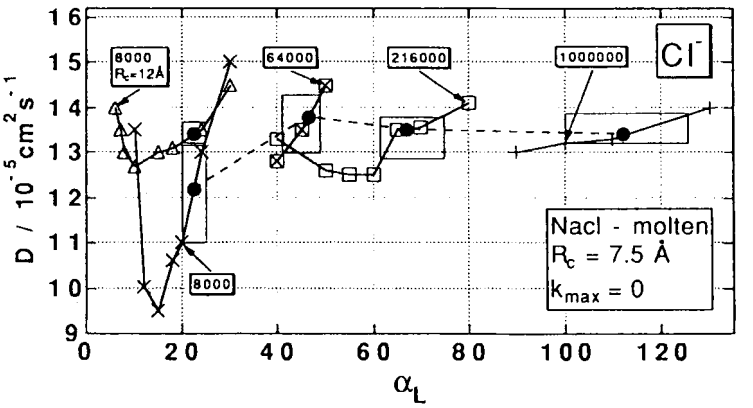

Figure 10 Diffusion co-efficient of *Cl* ions for *NaCl* systems of different sizes (q-space term switched off). For comparison the curve with $R_c = 12$ Å for $N = 8000$ system is given. Other symbols as in Figure 9.

rectangles around them show the area corresponding to the ±0.5% range of the energy plateau. The circles have been joined by dashed lines. These lines then represent the values of the calculated quantities for optimal values of the α_L parameter (i.e. α_L corresponding to the energy plateau region). For smaller systems the quantities are very sensitive to the choice of α_L. For bigger systems this sensitivity becomes smaller. Nevertheless, for all systems, with optimal values α_L, these results can be regarded as acceptable, especially if one takes into account that their accuracy in an MD calculation is typically of the order of 10%.

4.2 Results for α-AgI

Figure 13 shows coulombic energy Φ^c vs α_L for two sizes of *α-AgI* system ($N =$ 2048 and $N = 108\,000$) and for different values of R_c and k_{max}. Figure 14 shows Φ^q term for two selected curves from Figure 13. The behaviour of Φ^c around the

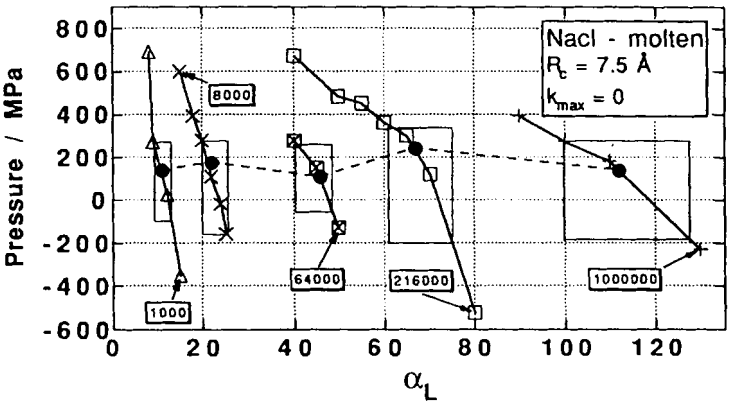

Figure 11 Internal pressure versus α_L for *NaC* systems of different sizes (q-space term switched off). For details see description in Figure 9.

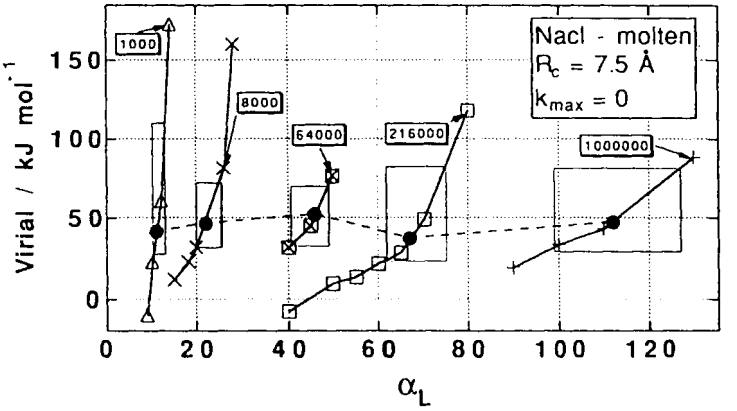

Figure 12 Virial versus α_L for *NaCl* systems of different sizes (q-space term swiched off). For details see description in Figure 9.

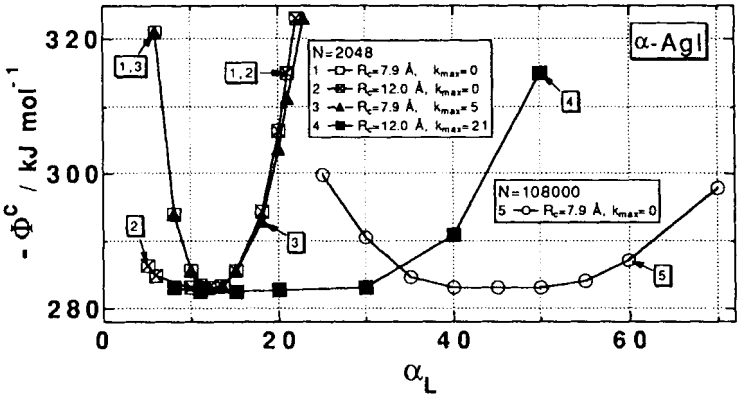

Figure 13 Coulombic energy of *α-AgI* systems for different values of cut-off radius R_c and different range of q-space summation k_{max}.

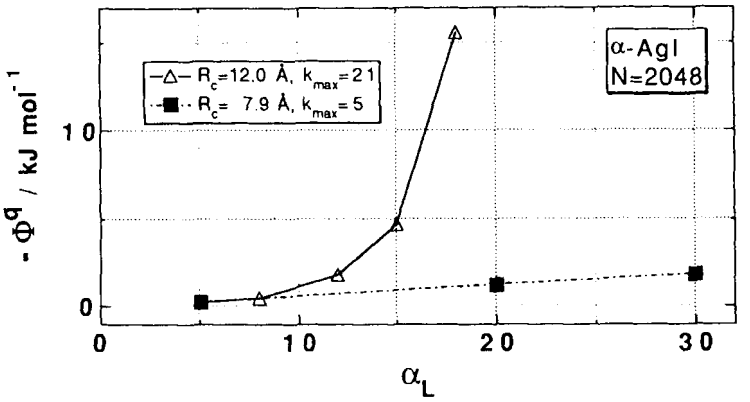

Figure 14 Energy q-space term of *α-AgI* system for different values of R_c and k_{max}.

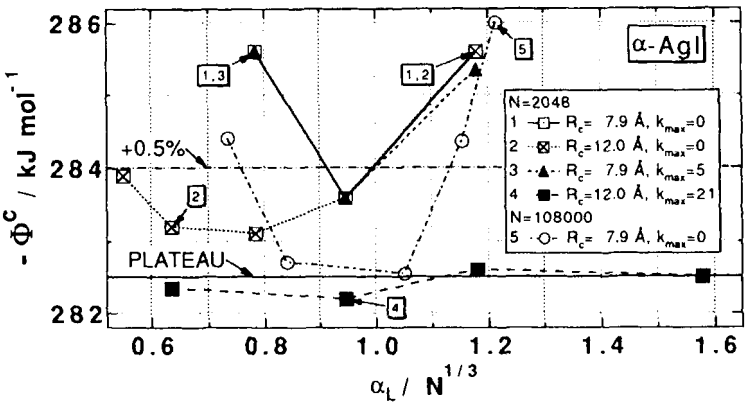

Figure 15 Coulombic energy for two *α-AgI* systems in the range of 0 to 0.5% around the plateau level.

$\alpha_L = N^{1/3}$ point is shown in Figure 15. Those figures are counterparts of Figures 2, 3 and 4 and their results support the conclusions drawn from the results for *NaCl* system. Also, the α_L vs $N^{1/3}$ dependence shown for *α-AgI* in Figure 16 is similar to that for *NaCl* (Figure 8). In Figure 17 is shown the diffusion coefficient for Ag^+ ions. The internal pressure of the system is shown in Figure 18 and the virial in Figure 19. The conditions of these simulations were analogous to those used for the *NaCl* system (i.e. $k_{max} = 0$ and a relatively small R_c; equal to 7.9 Å in this case). In Figures 18 and 19 the curves with $k_{max} = 21$ and bigger R_c have been included for comparison.

4.3 Results for δ-Bi_2O_3

The coulombic energy Φ^c vs α_L for various values of sizes N, k_{max}, and R_c is shown in Figures 20, 21 and 22. The Φ^q contribution is given in Figures 23, 24 and 25. The dependence of α_L on $N^{1/3}$ for $k_{max} = 0$ is presented in Figure 26.

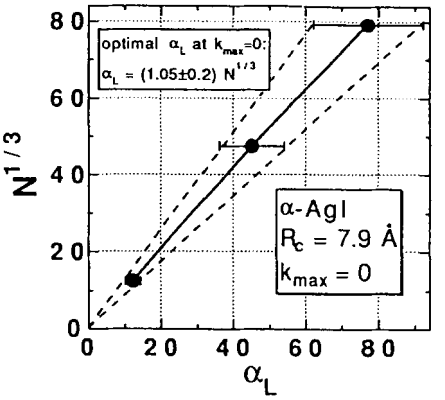

Figure 16 Dependency of the convergence parameter α_L on the system's size N. The closed circles correspond to the energy plateau while horizontal bars to the 0.5% region. Dashed lines show estimted range of the optimal value of α_L.

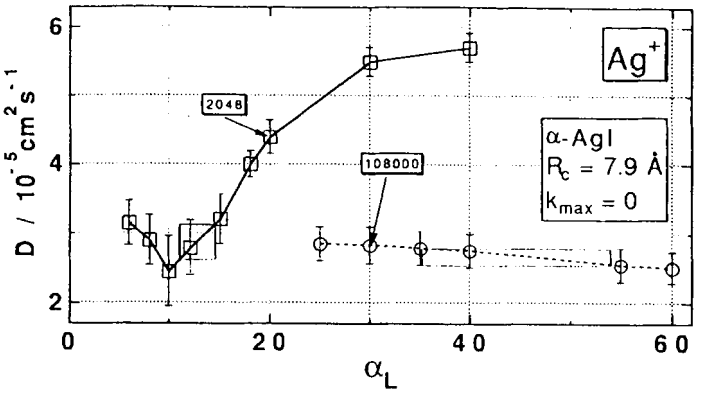

Figure 17 Diffusion coefficient of *Ag* ions in α-AgI system. Rectangulars denote 0 to 0.5% range from the plateau level.

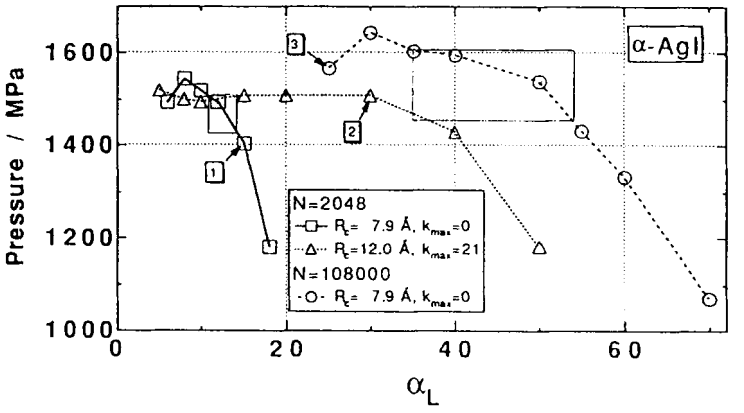

Figure 18 Internal pressure versus α_L in α-*AgI* system. Rectangles for $k_{max}=0$ denote 0 to 0.5% range from plateau level.

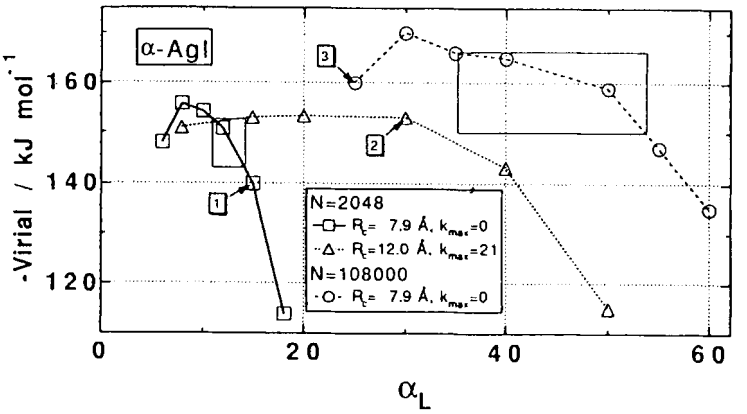

Figure 19 Virial versus α_L in *α-AgI* system. Rectangles for $k_{max} = 0$ denote 0 to 0.5% range from plateau level.

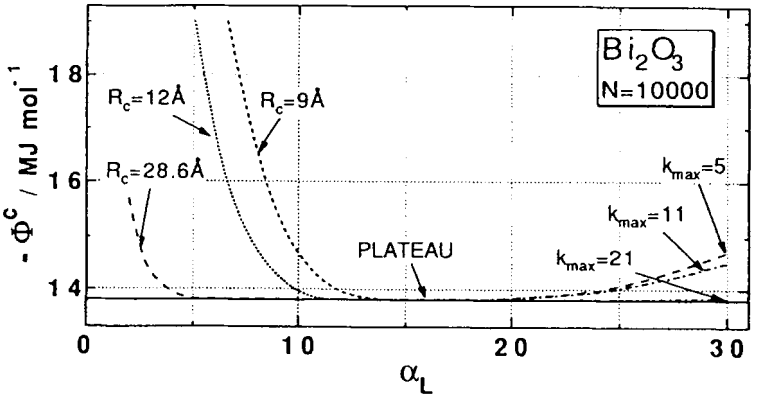

Figure 20 Coulombic energy of $N = 10\,000$ Bi_2O_3 system for different values of cut-off radius R_c and different range of q-space summation k_{max} (cf. Figure 1).

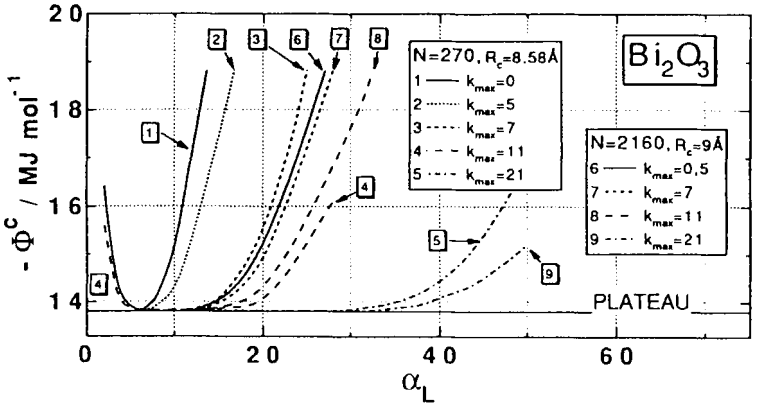

Figure 21 Coulombic energy of two smaller Bi_2O_3 systems for different values of k_{max} and $R_c = 9$ Å (for $N = 270$ system the maximum value available is $R_c = 8.58$ Å).

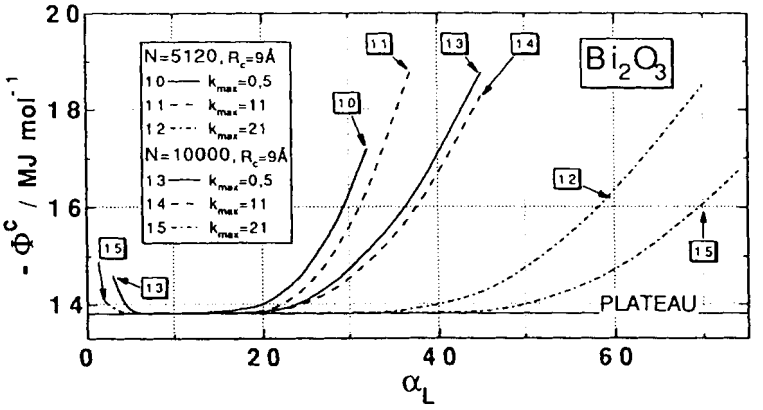

Figure 22 Coulombic energy of two larger Bi_2O_3 systems for different values of k_{max} and $R_c = 9$ Å.

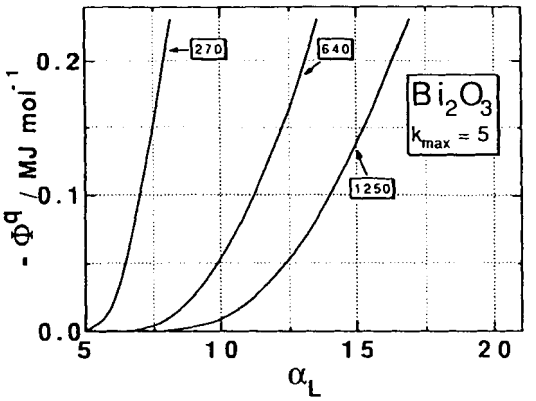

Figure 23 Energy q-space term of Bi_2O_3 system of different sizes with $k_{max} = 5$ and $R_c = 9$ Å ($R_c = 8.58$ Å for $N = 270$).

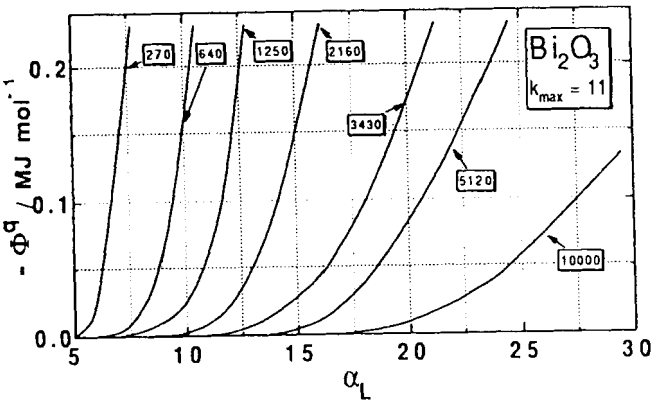

Figure 24 Energy q-space term of Bi_2O_3 system of different sizes with $k_{max} = 11$ and $R_c = 9$ Å ($R_c = 8.58$ Å for $N = 270$).

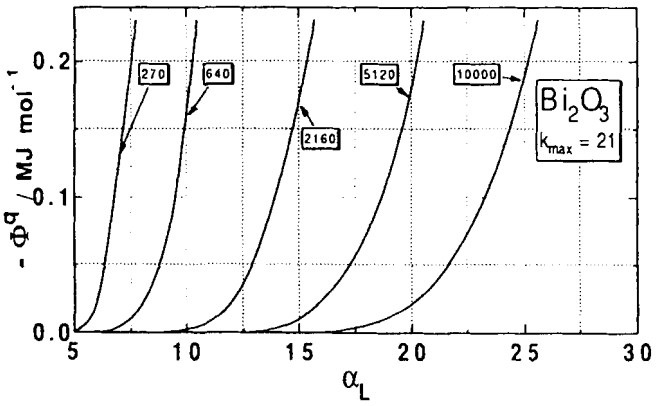

Figure 25 Energy q-space term of Bi_2O_3 system of different sizes with $k_{max} = 21$ and $R_c = 9$ Å ($R_c = 8.58$ Å for $N = 270$).

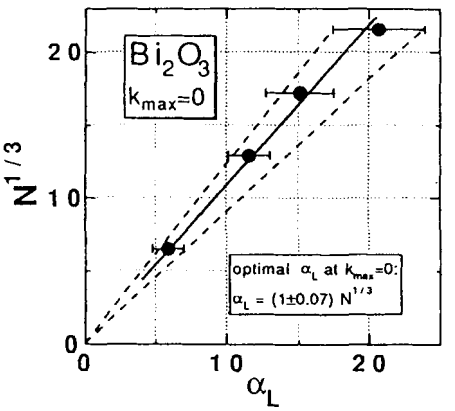

Figure 26 Dependence of the convergence parameter α_L on the system's size N. The closed circles correspond to the energy plateau while horizontal bars to the ±0.5% region. Dashed lines show estimated range of the optimal value of α_L.

4.4 Results for $SrCl_2$

Results for $SrCl_2$ are in general agreement with those for *NaCl*, *α-AgI* and *δ-Bi_2O_3* and they support the conclusions expressed in previous sections. For this reason and to avoid the inclusion of a large amount of data, these results are not shown here in detail.

5 DISCUSSION AND CONCLUSIONS

The results presented show that for large enough systems and properly chosen values of the α_L and R_{CD} parameters the q-space term can be omitted in the simulation of ionic systems. Its omission results in a significant reduction in CPU time and computer memory. In Figure 27 is shown the dependence of $\alpha_L \times R_{CD}$ on $N^{1/3}$ for $k_{max} = 0$ for the smallest values of R_{CD} that still hold the system in the

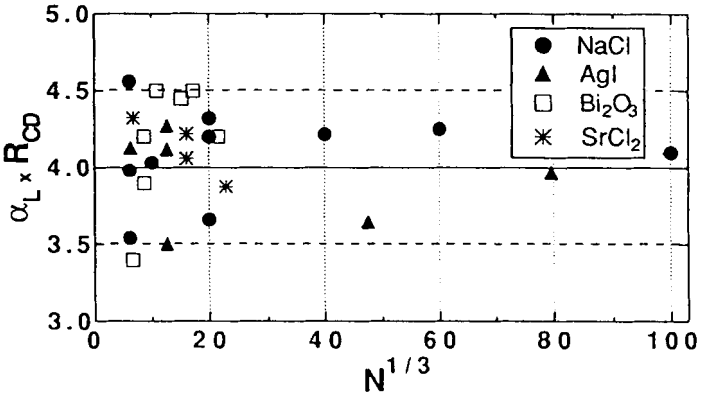

Figure 27 Dependence of the product of the dimensionless cut-off radius R_{CD} and α_L on the size N of the ionic systems examined.

plateau region. The values of $\alpha_L \times R_{CD}$ are approximately constant for all systems examined. These values are close to 4 and more appropriately they are given by equation 18 (α_1 parameter). The highest speed performance may be obtained by imposing small values of R_{CD} and $\alpha_L = \alpha_L^l = \alpha_L^u$ (equations 15 and 16). In practice, however, it is better to use more secure conditions for a simulation. For $k_{max} = 0$ these conditions should be chosen as follows:

(a) R_{CD} somewhat bigger than $\alpha_1 \cdot N^{-1/3}$ (20a)

(b) α_L somewhat smaller than $N^{1/3}$ (20b)

Some details on the choice of optimal parameters are given below.

Lower and upper limits of α_L. One of the conclusions of this work is the existence of well defined minimum for α_L and not existence a maximum value. Regardless of the system under consideration and conditions of simulations the used value of α_L cannot be smaller than α_1. The smallest value is approximately equal to 4 for all systems examined (equation 18). On the other hand an analogous restriction for upper limit of α_L does not exist; depending on the conditions of a simulation (size of the system N, cut-off radius R_c, range of q-space summation) the value of α_L can be really large.

Range of r-space summation. Recommended values of the cut-off radius R_c that have been found for the systems examined are shown in column 2 of Table 2. The

Table 2 Recommended values of cut-off radius, number of nearest neighbours and minimum size of the systems (for more details see text).

System	R_c/Å	$R_{CD} \times N^{1/3}$	NN_{av}	N_{min}
NaCl	7.5	4.7	56	600
α-AgI	7.9	5.0	64	700
δ-Bi_2O_3	8.8	6.6	153	1500
$SrCl_2$	8.5	5.4	83	800

corresponding values of $R_{CD} \cdot N^{1/3}$ are shown in column 3. Both these values are constants for a given system (they do not depend on N). The second value ($R_{CD} \cdot N^{1/3}$) is simply related to the average number of nearest neighbours by $NN_{av} = \pi \cdot R_{CD}^3 \cdot N/6$. Therefore the range of r-space summations can be expressed, in a convenient manner, in terms of NN_{av}. The values of NN_{av} given in column 4 in Table 2 represent a minimum recommended number of nearest neighbours that needs to be included in the calculation of the r-space forces to get accurate values of coulombic forces and energy. The values for the systems examined range from 56 to 153 and they are roughly 2–4 times bigger than those usually required for the calculation of short-range forces (the latter are order of 30–40). This point can be simply expressed as follows: *in order to exclude the q-space term it is necessary to increase the range of the r-space summation* (over the range which is adequate for short-range forces) *and adjust the value of* α_L *to make this summation rapidly convergent.*

The amount of CPU time required for calculation of the r-space term for a certain value of NN_{av} (equation 8) corresponds to the amount of time required for the calculation of q-space term with a certain value of k_{max} (equation 10). This relation is shown in Figure 28 where an equivalent number of nearest neighbours for each value of k_{max} has been given in brackets. Although CPU times in Figure 28 are given in msec (on the Hitac M-880) in fact they can be considered as given in arbitrary units since the relation between equations (8) and (10) is expected to be more or less the same on other serial computers. As can be seen, including the q-space summation with $k_{max} = 21$ requires the same amount of CPU time as the calculation in r-space over 16 300 nearest neighbours. The required value of NN_{av} depends on the system under consideration. If therefore it happens that in a system with strong coulombic interactions this number becomes very large, it might be more profitable to apply a conventional approach, i.e. use a smaller R_c and include the q-space term.

Minimum size of the system. The minimum size of the system (N_{min}) for which the q-space term may be neglected can be related to the NN_{av} value. In order to weaken the influence of self-correlations in the system the side length of MD box L

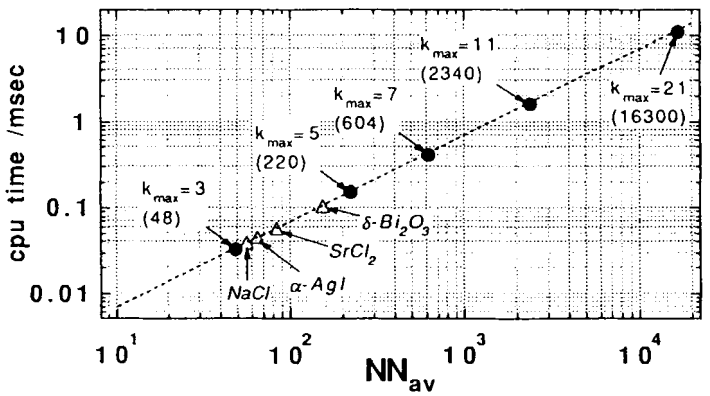

Figure 28 CPU time for the r-space term calculation as a function of NN_{av}. The closed circles show the equivalent cpu time for the q-space term calculation. For each value of k_{max} are given the values of NN_{av} (in brackets) that require the same amount of cpu time.

should be bigger than the required cut-off radius. This minimum size can be therefore expressed as,

$$N_{min} = c_1 \cdot NN_{av} \tag{21}$$

where $c_1 = 6c_O^3/\pi$ with $c_O = L_{min}/(2R_c)$. The bigger c_O the smaller the correlations; however a reasonable choice is to apply c_O between 1.5 to 2.0, which corresponds to c_1 between 6.5 to 15.3. Roughly, one may assume a value of c_1 of the order of 10 (i.e. L_{min} bigger by 75% than $2R_c$). The values of N_{min} estimated in such a way are given in the last column in Table 2. In fact those values should be considered as rough estimates only. There is not any specific minimum for which the relations (20) are valid (e.g. the use of larger R_c allows one to reduce a minimum size of system). On the other hand it is easier to hold a bigger system in the plateau region since its range increases with the value of N. For the systems smaller than 10^3 particles and the values of α_L and R_{CD} given by equation (20), the plateau region becomes very narrow and therefore it is necessary to choose them with great precision to ensure that the system is still within the plateau region. In fact, since the relations (20) have been found through MD simulations they should be considered rather as preliminary choices for a substance other than the four examined here: therefore it may be necessary to adjust the values of R_{CD} and α_L for a specific system.

Efficiency of simulation. In Figure 29 is shown the dependence of CPU time on the size of a system. Curves 1a, b, c and d are given for the method presented and for $k_{max} = 0$, 5, 11 and 21, respectively. These times are given for the Hitac M-880 mainframe and they are compared with other methods. Curve 2 represents CPU time of the $O(N^{3/2})$ PPL method (9) on the Fujitsu VP 200 supercomputer, while curve 3 is the CPU time for a classical $O(N^2)$ method [15] performed on the ETA 10P supercomputer. This comparison should be viewed with some tolerance since different computers are involved (however, the speeds of these computers are estimated to be within a factor of 2).

A *practical limit* marked in the Figure 29 has been assumed to be 50 CPU seconds per time step (a thousand time steps of such a simulation would require 14 CPU

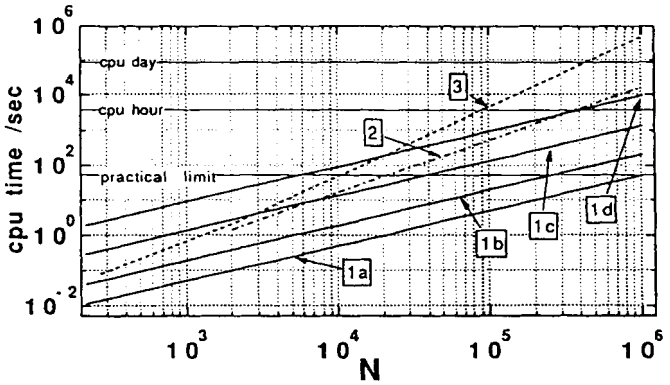

Figure 29 CPU time per simulation step for different MD methods. Symbol 1 denotes the $O(N)$ method presented here (1a, 1b, 1c and 1d correspond to $k_{max} = 0$, 5, 11 and 21 respectively), symbol 2 the PPL method [9] and symbol 3 a classical MD method [15]. The cpu times are given for different computers (for more details see text).

Table 3 Acceleration of the method presented (1a,c) and the PPL method (2) over a classical $0(N^2)$ method (3) (data from Figure 29).

Method	$N = 10^4$	$N = 10^6$
1a	87	8700
1c	3.3	330
2	3	30
3	1	1

hours). According to the data shown a practical limit for a classical $0(N^2)$ method is order of $N = 10^4$ particles on the ETA 10P supercomputer (or less on a slower computer). For the PPL method this limit is order of $N = 2 \times 10^4$ particles on the Fujitsu VP 200 supercomputer, while for the method presented it is of the order of $N = 10^6$ particles (on a serial computer). For $N = 10^6$ the method presented is faster than the other two by orders of magnitude (Table 3). Omission of the q-space term accelerates the simulation by one to two orders of magnitude depending on the value of k_{max} used (by 4 for $k_{max} = 5$, 25 for $k_{max} = 11$ and 180 for $k_{max} = 21$).

6 SUMMARY

An efficient order of N method for the simulation of condensed matter systems has been presented. The short-range interactions are calculated by the method based on the scalar pyramid algorithm [6, 8] while the long-range interactions are calculated by the Ewald method. The only possibility of further and substantial acceleration of an $0(N)$ MD simulation is to exclude the q-space term in the Ewald summation. It has been shown that such an exclusion is possible provided that suitable parameters are appropriately chosen. Omission of the q-space term results in substantial reductions of both CPU time and computer memory. With this method, due to its high speed performance, it is possible to simulate on a mainframe systems containing of the order of 10^6 particles, which is far more than is possible for any other method. On a general purpose microcomputer or workstation it can be used to simulate systems of 10^3 to 10^4 particles.

References

[1] P.P. Ewald, "Die Berechnung optischer und elektrostatischer Gitterpotentiale", *Ann. Physik*, **64**, 253 (1921).

[2] J. Boris, "A vectorized near neighbours algorithm of order N using a Monotonic Logical Grid", *J. Comput. Phys.*, **66**, 1 (1986).

[3] D.C. Rapaport, "Large-Scale Molecular Dynamics Simulation Using Vector and Parallel Computers", *Comput. Phys. Reports*, **9**, 1 (1988).

[4] G.S. Grest, B. Dunweg and K. Kremer, "Vectorized Link Cell Fortran Code for Molecular Dynamics Simulation for a Large Number of Particles", *Computer Phys. Commun.*, **55**, 269 (1989).

[5] Z.A. Rycerz and P.W.M. Jacobs, "Molecular dynamics simulation program of order N for condensed matter. I MDPYRS1: scalar pyramid, short-range interactions", *Comput. Phys. Commun.*, **60**, 53 (1990).

[6] Z.A. Rycerz, "Molecular dynamics simulation program of order N for condensed matter systems.

II MDSPNL: pyramid with neighbour list, short-range interactions", *Comput. Phys. Commun.*, **61**, 361 (1990).
[7] Z.A. Rycerz and P.W.M. Jacobs, "A Vectorized Program of Order N for Molecular Dynamic Simulation of Condensed Matter, II. MDSLAB1: Slab, Short-Range Interactions", *Computer Phys. Commun.*, **62**, 145 (1991).
[8] Z.A. Rycerz and S. Ihara, "Molecular dynamics simulation of multi-million particle systems", to be published.
[9] J.W. Perram, H.G. Petersen and S.W. De Leeuw, "An algorithm for the simulation of condensed matter which grows as the 3/2 power of the number of particles", *Molecular Phys.*, **65**, 875 (1988).
[10] Z.A. Rycerz and P.W.M. Jacobs, "Ewald summation in the molecular dynamics simulation of large ionic systems: The cohesive energy", *Molecular Simulation*, **8**, 197 (1992).
[11] D.J. Adams and G.S. Dubey, "Ewald sum made easy", *IQCSCP*, **22**, 39 (1986).
[12] D.J. Adams and G.S. Dubey, "Taming the Ewald sum in the computer simulation of charged systems", *J. Comput. Phys.*, **72**, 156 (1987).
[13] S. Kuwajima and A. Warshel, "The extended Ewald method: a general treatment of long-range electrostatic interactions in microscopic simulations", *J. Chem. Phys.*, **89**, 3751 (1988).
[14] L. Greengard and V. Rokhlin, "On the evaluation of electrostatic interactions in molecular modeling", *Chemica Scripta*, **29A**, 123 (1989).
[15] N. Anastasiou and D.Fincham, "Programs for the dynamic simulation of liquids and solids. II MDIONS: rigid ions using the Ewald sum", *Comput. Phys. Commun.*, **25**, 159 (1982).
[16] M.J.L. Sangster and M.Dixon, "Interionic potentials in alkali halides", *Adv. Phys.*, **25**, 247 (1976).
[17] W. Smith, "Fortran code for the Ewald summation method", IQCSCP, **21**, 37 (1986).
[18] Z.A. Rycerz and S. Ihara, "MULISP program", Int. Rep., Hitachi Ltd. (1991).
[19] P.W.M. Jacobs and Z.A. Rycerz, "Molecular dynamics of α-*AgI*", *Cryst. Latt. Def. and Amorph. Mat.*, **15**, 337 (1987).
[20] P.W.M. Jacobs and Z.A. Rycerz, unpublished work.
[21] D.A. MacDonaill, P.W.M. Jacobs and Z.A. Rycerz, "Molecular dynamics simulation of the fast ion conductor δ-Bi_2O_3", *Molecular Simulation*, **3**, 155 (1989).
[22] J. Moscinski and P.W.M. Jacobs, "Computer simulation of defect motion in model normal and 'fast–ion' conductors. I, *NaCl* and *AgCl*", *Proc. Roy. Soc. London*, **A 398**, 141 (1985).

Note: The IQCSCP abbreviation denotes – *Information Quarterly for Computer Simulation of Condensed Phases*, CCP5, SERC, Daresbury Laboratory, Daresbury, England.