John von Neumann Institute for Computing

A Fast Wavelet Based Evaluation of Coulomb Potentials in Molecular Systems

Godehard Sutmann, Silke Wädow

published in

NIC Workshop 2006,
From Computational Biophysics to Systems Biology,
Jan Meinke, Olav Zimmermann,
Sandipan Mohanty, Ulrich H.E. Hansmann (Editors)
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **34**, ISBN-10: 3-9810843-0-6,
ISBN-13: 978-3-9810843-0-6, pp. 185-188 , 2006.

http://www.fz-juelich.de/nic-series/volume34

A Fast Wavelet Based Evaluation of Coulomb Potentials in Molecular Systems

Godehard Sutmann and Silke Wädow[†]

Central Institute for Applied Mathematics and
John von Neumann Institute for Computing,
Research Centre Jülich, 52425 Jülich, Germany
E-mail: g.sutmann@fz-juelich.de

1 Introduction

Long range interactions often play an important role in complex molecular systems. Due to charged patterns or dipolar structures, molecules interact via Coulomb interactions, thereby influencing dynamical processes or static structures. Modeling charged or polar systems on a microscopic level is often limited by the number of system constituents, since calculation of long range interactions are computationally very demanding, as they scale like $\mathcal{O}(N^2)$ since all interactions between particle pairs have to be considered explicitely. However, approximating infinitely large systems by relatively small ones and imposing periodic boundary conditions at the same time, results in evaluating lattice sums, which usually converge conditionally. Splitting of lattice sums and applying certain controllable approximations leads to handy expressions which, however, still scale like $\mathcal{O}(N^{3/2})$.

Due to this computational burden, research was focused on developing methods, which overcome the quadratic scaling and reduce it to an optimal $\mathcal{O}(N)$, or at least an $\mathcal{O}(N \log N)$ behavior. In general the proposed methods can be classified into mesh-free and mesh-based algorithms, both existing for periodic and open system geometries. Mesh-free algorithms, like the Fast Multipole Method (FMM)[1], should in principle be more promising, since the discretization error is avoided. The problem encountered here is however a very evolved and demanding work to implement the method properly. Alternatives may be found in mesh-based methods, where fast methods like multigrid techniques[2,3] may be applied. All these methods, scaling like $\mathcal{O}(N)$ have in common that they work in a hierarchical way, i.e. collecting different contributions on different length scales or levels of resolution.

In the present article a new mesh-based method is proposed, which makes use of a fast Wavelet transform technique, thereby reducing the computational complexity to $\mathcal{O}(N)$.

2 Calculation of Potential Energies in Particle Systems

The potential at a charges' position with index i is given by

$$\phi(\mathbf{r}_i) = \sum_{i \neq j}^{N} \frac{q_j}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1}$$

[†]Present address: INFORM GmbH, Pascalstr.23, D-52076 Aachen, Germany

where N is the number of charges in the system, q_j the charge of particle j and $\mathbf{r}_j$ its position. The interaction energy of particle i with all other particles is then simply given by

$$U_i(\mathbf{r}_i) = q_i \phi(\mathbf{r}_i) \tag{2}$$

Eq. 1 may also be expressed in short hand notation as matrix-vector product

$$\mathbf{\Phi}(\{\mathbf{R}\}) = \mathbf{A}(\{\mathbf{R}\})\mathbf{Q} \tag{3}$$

where $\mathbf{\Phi} = \{\phi_1, \ldots, \phi_N\}$, $\mathbf{Q} = \{q_1, \ldots, q_N\}$ and matrix elements are given by $A_{ij} = 1/|\mathbf{r}_i - \mathbf{r}_j|$. The interaction energy could then simply be written as $\mathbf{U} = \mathrm{diag}(\mathbf{Q}\mathbf{\Phi}^T)$. The calculation of the potential via Eq. 3 uses a constant charge vector, but the matrix elements change from step to step if the particles move along their trajectories. Therefore this approach would imply to recalculate matrix elements, which is an $\mathcal{O}(N^2)$ operation. Therefore it would be desirable to shift the time dependence from the matrix to the vector, which would imply only an $\mathcal{O}(N)$ operation. This may be achieved by introducing a mesh, onto which the particle charges are distributed and which has a constant grid spacing throughout the time evolution of the system (even this constraint may be relaxed by rescaling of lengths). Therefore, the grid based summation would be

$$\hat{\mathbf{\Phi}}(\{\hat{\mathbf{R}}\}) = \hat{\mathbf{A}}(\{\hat{\mathbf{R}}\})\hat{\mathbf{Q}} \tag{4}$$

where '$\hat{\cdot}$' means mesh based quantities. Nevertheless, also with this modification, the complexity is still $\mathcal{O}(N^2)$ due to a dense matrix-vector product. To obtain an efficiency gain, it will therefore be necessary to transform the matrix into a sparse representation.

3 Principles of Wavelet Based Summations

In principle, wavelets are a tool to analyse data on different time and length scales[4]. In contrast to Fourier analysis, signals are developed in a basis, which have local support. Wavelet analysis is one class of multiscale analysis methods, which splits signals into fine and coarse contributions in a hierarchical scheme. If the original signal f is within a subspace V_{-1} of $L^2(\mathbb{R})$, then high- and low-frequency parts may be projected out via

$$f = \mathcal{P}_0 f + \mathcal{Q}_0 f \tag{5}$$

where V_{-1} is now the sum of orthogonal subspaces, containing low- (V_0) and high-frequency (W_0) parts

$$f \in V_{-1} = V_0 \oplus W_0 \tag{6}$$

This procedure may be continued up to a certain level M, where the subspaces W_i are subject to further splittings into high- and low-frequency contributions, thereby leading to the splitting scheme

$$f = \mathcal{P}_M f + \sum_{k=0}^{M} \mathcal{Q}_k f \tag{7}$$

Wavelet analysis works as a combined low- and high-pass filter, thereby generating coefficients, describing the strength of signal change on a given coarsening level. Small coefficients thus contain only little information to details on a given level. In image processing it

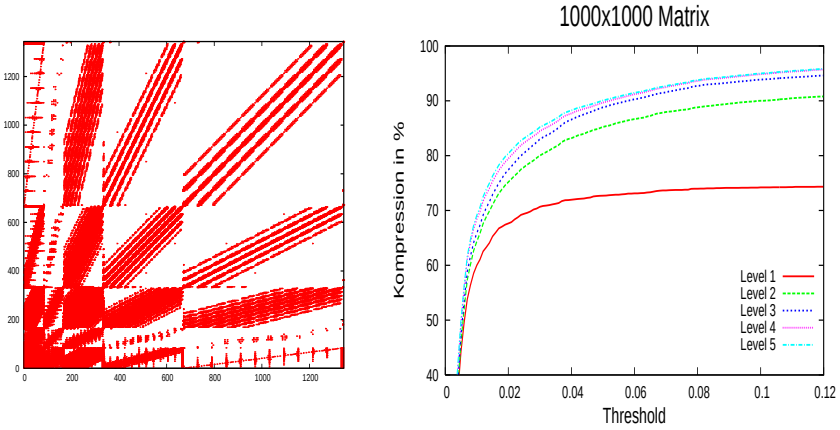

Figure 1. Left: Structure of the inverse distance matrix in wavelet space after thresholding. Right: Compression of original matrix as a function of threshold value x for different levels M of multiresolution analysis (cmp. Eq. 7). In both cases a Daubechies-12 wavelet basis was used[5].

has been observed that after wavelet transformation of pixel distributions of images, most coefficients are very small. Introducing a threshold value, below of which the absolute value of coefficients is set to zero, the wavelet representation of pictures become sparse. A back-transform using the sparse matrix representation contains still the essential features of the picture.

This principle may also be applied to other matrices, where the number of small coefficients may depend on the symmetry of the matrix. In the present case it is applied to the inverse-distance matrix $\mathbf{A}$. Since the diagonal of that matrix is on one hand formally infinity, but on the other hand contains contributions to the self-energy of the particle, it is not of interest and therefore is simply set to zero. If charges are distributed to the grid by the nearest grid point method[6], no further corrections to self-energies have to be applied. It is, nevertheless, stressed that this approximation induces relatively large discritization errors.

A wavelet transform may be represented by an orthogonal matrix $\mathcal{W}$, which leads in the present case to the representation

$$\mathcal{W}\mathbf{\Phi} = \mathcal{W}\mathbf{A}\mathcal{W}^T\,\mathcal{W}\mathbf{Q} \tag{8}$$

which can be written as

$$\tilde{\mathbf{\Phi}} = \tilde{\mathbf{A}}\tilde{\mathbf{Q}} \tag{9}$$

where '$\tilde{\cdot}$' means a wavelet transformed quantity. Applying a threshold operation $\mathcal{T}_x$ to $\tilde{\mathbf{A}}$, where x is the threshold size, transforms $\tilde{\mathbf{A}}$ into a sparse representation (cmp. Fig. 1). Since the distance between grid points in the geometrical domain is kept constant during time, the wavelet transform of $\hat{\mathbf{A}}$ has only to be performed in the beginning of the simulation. The time dependence is completely shifted to the vector $\hat{\mathbf{Q}}$, thereby reducing the transformation to a N-dimensional quantity. Applying a fast wavelet transform the computational complexity is reduced to $\mathcal{O}(N)$. Finally, to obtain the potential on every grid

point, a fast backtransform of $\mathcal{O}(N)$ complexity is applied in order to obtain $\hat{\mathbf{\Phi}}$. The potential at the charges' position is then obtained via interpolation from the grid points, which may be performed via polynomial approximation of the potential surface. In numerical tests, it was found that the sparse matrix-vector product may also be performed with $\mathcal{O}(N)$ complexity while keeping the approximation error constant. The latter, of course, depends on the size of the threshold x, but it may be bound to relative errors of 10^{-2} and below.

4 Conclusions

A fast method of order $\mathcal{O}(N)$ for the calculation of electrostatic potentials has been described. It is based on a wavelet transformation technique which, in combination with a threshold operation, transforms the inverse distance matrix, $\mathbf{A}$, into a sparse representation, thereby dramatically reducing the computational time of matrix-vector multiplication. The proposed method may be extended to calculate interaction energies and forces between particles. To increase the accuracy, other schemes may be applied to smear charges to the grid and to interpolate the potential back. Here it is necessary to treat self energies and near-field contributions in a proper way. Work in this direction is in progress.

References

1. L. Greengard. *The rapid evaluation of potential fields in particle systems*. MIT press, Cambridge, 1988.
2. R. D. Skeel, I. Tezcan, and D. J.Hardy. Multiple grid methods for classical molecular dynamics. *J. Comp. Chem.*, 23:673–684, 2002.
3. G. Sutmann and B. Steffen. A particle-particle particle-multigrid method for long-range interactions in molecular simulations. *Comp. Phys. Comm.*, 169:343–346, 2005.
4. G. Beylkin. Wavelets, Multiresolution Analysis and Fast Numerical Algorithms, 1991. A draft of INRIA Lecture Notes.
5. I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, 1992.
6. R. W. Hockney and J. W. Eastwood. *Computer simulation using particles*. McGraw-Hill, New York, 1981.