FINAL REPORT

DOE DE-FG02-91ER25099

Numerical Methods for Molecular Dynamics

Robert D. Skeel, University of Illinois

DOE/ER/25099--T1

Amount: $75,000 total
Period: August 1, 1993 to January 31, 1996.

Work through July 31, 1993 is covered in previous annual reports, so only work done since then is described here. This award was joint with one from the NSF Applied Mathematics Program having the same title:

Award number: NSF DMS 93-04268
Amount: $61,500 plus REU supplement of $7,500
Period: July 15, 1993 to June 30, 1996
The work reported here was supported by both grants.

The aim of this research is to explore ideas for more efficient numerical integrators for molecular dynamics (MD) [1] and where needed to develop appropriate theoretical tools. Emphasis is on macromolecules [9] and techniques suitable for implementation in the biomolecular dynamics program NAMD [11]. Listed below are the main accomplishments during the past two or so years. First are listed algorithmic developments suitable for implementation:

1. A rule was developed and supported by rigorous analysis for the assignment of stepsizes to differing interactions in a multiple time stepping scheme.
2. A very cheap scheme for postprocessing of trajectories was developed that removes most of the high frequency error from the trajectory without sacrificing the ability to use energy conservation as a meaningful error indicator.
3. Fourth-order integration schemes have been found, which are more efficient than existing fourth-order schemes.
4. Variations of implicit schemes have been found that for the most part retain their desirable properties but are computationally much less costly.
5. Algorithms were found for MD with holonomic constraints that are considerably faster than the very popular SHAKE iteration.

Second are listed more theoretical developments helpful for further exploration of new algorithms.

6. A well known symplectic multiple timestepping method has been proved to be unsound for long timesteps, but an enhancement has been found, which has been proved to be symplectic and to retain its order of accuracy for long timesteps.
7. An analytical tool has been developed based on elementary Hamiltonians which simplifies the search for symplectic methods of higher *effective* accuracy.
8. Methodology has been demonstrated for finding formulas of optimal stability, and these have been shown also to be high in accuracy.
9. The possibility of maintaining symplectic integration in floating-point arithmetic has been demonstrated.
10. Symplectic integration has been shown to be strongly linked to conservation of angular momentum.

The remainder of this section describes this work in greater detail.

Empirical evidence (e.g., [16, 13]) suggests that to get better qualitative behavior and slower error growth one should use integrators that are symplectic and/or time reversible. Theoretical considerations favor symplectic integrators because they produce numerical solutions that are very nearly the exact

solution of a slightly different Hamiltonian system [16, p. 132]. This is especially important for MD where there are concerns about the meaningfulness of long-time integrations due to the exponential growth (on a picosecond time scale) of errors in particle trajectories [1, 7]. The potential energy functions ("force fields") of MD are empirical approximations to the correct quantum mechanical description, but presumably the statistical quantities that are being computed are not too sensitive to these errors. Thus it seems to desirable that temporal discretization errors are not only modest in size but have a quality similar to errors in empirical potential energy functions. The very popular Verlet method is symplectic and it has recently been shown [2] that Verlet-SHAKE also computes position trajectories which are those of a symplectic method.

Most of the interactions in MD are weak and slowly varying and could be resolved with a stepsize that is much greater than the stepsize used for the much smaller number of rapidly varying interactions. This feature of MD can be exploited by *multiple time step* (MTS) methods, in which different terms of the right-hand side of the differential equation are evaluated at different time increments Δt. Most MTS extensions of the Verlet method [8] are nonsymplectic, but a symplectic "impulse" MTS method (Verlet-I, RESPA) has been proposed by my collaborators [6] and independently by a group at Columbia [19]. One question of importance is how to assign stepsizes to different interactions. A practical answer for bonded interactions, supported by rigorous analysis, is given in [LiSZ9x]. We have also studied this question at length for electrostatics interactions. A tentative result from this work is an estimate of $O(N^{4/3})$ for the computational complexity of MTS compared to $O(N^2)$ if a direct calculation of the electrostatics forces is done every timestep. Fast electrostatic methods can do the calculation in $O(N)$ or $O(N \log N)$ operations but with a large "hidden" constant.

In its basic form, MTS is appropriate only for bonded interactions in which a suitable fixed stepsize (varying from one interaction to another) is used for each interaction. For nonbonded interactions it is very important to be able to vary the stepsize as the distance between particles changes. A practical way to vary the stepsize, proposed in [14], is to split the potential artificially into a sum of potentials of decreasing range but also decreasing smoothness. Implementation details are being developed for a highly efficient second-order full multiple timestep scheme, and numerical experiments are being conducted. In principal the splitting consists of an infinite series of potentials, thus allowing arbitrarily small stepsizes in very hard interactions such as Lennard-Jones collisions.

It has been known for a while that the symplectic MTS scheme is susceptible to resonance [6, 8] and one year ago it became apparent that for long timesteps this is a grave problem: we have shown exponential growth as $t \to \infty$, Δt fixed, for linear problem. Also the scheme may not be consistent—it has not yet been determined whether or not the method suffers an *order reduction* to zero. This poor behavior agrees with experiments of Grubmüller [7]. We have made significant progress in finding a more useful symplectic MTS method.

It has been observed by various authors, e. g., [15, 20] in the context of Hamiltonian systems, that the accuracy of an integration scheme depends on the interpretation of the numerical solution. We discovered this independently in efforts [LiSZ9x] to determine an optimal choice of timestep ratios for multiple timesteps. More specifically, the idea is that we associate with a method a transformation that relates the correct values to their slightly encoded counterparts used by the algorithm. This idea opens up new possibilities for symplectic integration. A first systematic look was undertaken in [10] for Runge–Kutta-like methods with the idea that the transformation should have a practical implementation as a Runge-Kutta "corrector." (We use quotes here to indicate that this is quite different from the corrector in a predictor-corrector method.) However, one may not wish to actually perform this transformation, in particular, for MD where initial velocities are chosen at random. The prime motivation for considering transformations is to able to identify more reliably those methods that most accurately reproduce the long-time dynamics. In our work [LoSS97] we have simplified the algebra by instead using flows of sums

of *elementary Hamiltonians* to represent the pre- and post-processing transformations. Nonetheless, there may be reason to actually perform the processing (or at least the postprocessing): one might be interested in monitoring conservation of energy[1] or one might be interested in computing quantities whose accuracy is affected by the high frequency error that is present in an unprocessed trajectory. We have found [LoSS96a] a highly efficient postprocessing algorithm based on differencing which is nearly symplectic and whose cost is negligible.

Experience with numerical methods suggests that fourth-order methods are likely to be optimal for low-accuracy solution of problems like MD. Among methods that use a Hessian–vector product, we have found a simple 4th-order method [LoSS97] that is more efficient than the excellent scheme due to Rowlands [15]. (The cost of doing both a force and a Hessian–vector product is at most double the cost of evaluating the force alone.) Also studied anew were 3-stage schemes that use only force evaluations. In both cases an important tool was the idea of effective order. In the case of the 3-stage scheme another important consideration was the interval of stability: there is reason to think that a long stability interval might have the added benefit of yielding smaller error coefficients; for example, the 3-stage effectively fourth-order optimally stable scheme, which we found, has an error coefficient that is smaller than the well known Candy-Rozmus-Forest-Ruth [3, 4] method by a factor of one hundred!

Implicit methods have been proposed for MD in combination with Langevin dynamics [12] and in association with local linearizaton (normal mode analysis) [21]. The high cost of solving systems of nonlinear equations with dense Jacobian matrices has been made manageable by the use of an efficient minimizer TNPACK [17], which is able to exploit the special structure of macromolecular energy functions. To bring down the cost even more, we have found how to design (i) linearly implicit and (ii) mixed implicit-explicit methods that retain the property of being symplectic. This work is being written up.

In MD the stepsize is constrained by the small-amplitude high-frequency components of the motion that are due to bondlength stretching. For this reason it is popular to make the bondlengths rigid. This results in a system of differential equations with quadratic constraints. It has been popular to use the SHAKE iteration for solving the nonlinear algebraic equations that result. We have investigated [BKLS95] the use of Newton methods with some success. In addition, however, we have discovered that the use of overrelaxation with SHAKE leads to a great reduction in the number of iterations.

In view of the rapid exponential growth of perturbations to atomic trajectories, even the tiny errors due to roundoff and truncated multipole expansions should be viewed as potentially serious. It has been shown [18] that the effect of computing the trajectory of a symplectic integrator on a uniform lattice is equivalent to a very small perturbation of the Hamiltonian. We have shown how a lattice leapfrog method can be implemented in floating-point arithmetic at very little cost and how the analysis extends to include errors other than those due to the lattice.

It was shown [ZhSk95] that explicit symplectic Runge-Kutta-Nyström methods conserve angular momentum and that the result generalizes to multiple time steps and to constrained dynamics. The first two of these results also follow from a more abstract result of Ge [5].

1. List of publications acknowledging the DOE award.

[BKLS95] E. Barth, K. Kuczera, B. Leimkuhler, and R. D. Skeel, "Algorithms for constrained molecular dynamics," *J. Comput. Chem. 16*, 1192–1209, 1995.

[HKSS94a] M. Holst, R. Kozack, F. Saied, and S. Subramaniam, "Protein electrostatics: rapid multigrid-based Newton algorithm for solution of the full nonlinear Poisson-Boltzmann equation," *J. Biomol. Struct. Dyn. 11*, 1437–1445, 1994.

[HoSa93a] M. Holst and F. Saied, *J. Comp. Chem. 14*, 105–113, 1993.

[1] If both the time integrator and the processing are symplectic, then the fluctuations in total energy of the processed solution represents a sampling of the perturbation to the Hamiltonian due to the finite stepsize Δt.

[LeRS96] B. J. Leimkuhler, S. Reich, and R. D. Skeel, "Integration methods for molecular dynamics," in *Mathematical Approaches to Biomolecular Structure and Dynamics*, J. Mesirov, K. Schulten, and D. W. Sumners, eds., Springer-Verlag, tentatively 161–186, tentatively 1996.

[LiSZ9x] T. R. Littell, R. D. Skeel, and M. Zhang, "Error analysis of symplectic multiple time stepping," *SIAM J. Numer. Anal.*, to appear.

[LoSS96a] M. López-Marcos, J. M. Sanz-Serna, and R. D. Skeel, "Cheap enhancement of symplectic integrators," Proceedings, 1995 Dundee Conference on Numerical Analysis, to appear.

[LoSS96b] M. López-Marcos, J. M. Sanz-Serna, and R. D. Skeel, "Explicit symplectic integrators with maximal stability intervals," World Scientific Publishing Co., to appear in a book to celebrate Prof. A. R. Mitchell's 70th birthday.

[LoSS97] M. López-Marcos, J. M. Sanz-Serna, and R. D. Skeel, "Explicit symplectic integrators using Hessian–vector products," *SIAM J. Sci. Comput. 18*, 1997, to appear.

[Reic95] S. Reich, "Smoothed dynamics of highly oscillatory Hamiltonian systems," *Physica D 89*, 28–42, 1995.

[Reic96] S. Reich, "Enhancing energy conserving methods," *BIT Numer. Math.*, 1996, to appear.

[SkZS97] R. D. Skeel, G. Zhang, and T. Schlick, "A family of symplectic integrators: stability, accuracy, and molecular dynamics applications," *SIAM J. Sci. Comput. 18*, 1997, to appear.

[ZhSk95] M. Q. Zhang and R. D. Skeel, "Symplectic integrators and the conservation of angular momentum," *J. Comput. Chem. 16*, 365–369, 1995.

REFERENCES

[1] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids.* Clarendon Press, Oxford, 1987. Reprinted in paperback in 1989 with corrections.

[2] B. Leimkuhler and R. D. Skeel. Symplectic numerical integrators in constrained Hamiltonian systems. *J. Comput. Phys.*, 112(1):117–125, May 1994.

[3] J. Candy and W. Rozmus. A symplectic integration algorithm for separable Hamiltonian functions. *J. Comput. Phys.*, 92:230–256, 1991.

[4] E. Forest and R. D. Ruth. Fourth-order symplectic integration. *Physica D*, 43:105–117, 1990.

[5] Z. Ge. Equivariant symplectic difference schemes and generating functions. *Physica D*, 49:376–386, 1991.

[6] H. Grubmüller, H. Heller, A. Windemuth, and K. Schulten. Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions. *Molecular Simulation*, 6:121–142, 1991.

[7] H. Grubmüller and P. Tavan. Efficient algorithms for molecular dynamics simulations of proteins: How good are they? Manuscript, 1994.

[8] J. J. Biesiadecki and R. D. Skeel. Dangers of multiple-time-step methods. *J. Comput. Phys.*, 109:318–328, 1993.

[9] J. A. McCammon and S. C. Harvey. *Dynamics of Proteins and Nucleic Acids.* Cambridge Univ. Press, Cambridge, 1987.

[10] R. I. McLachlan. More on symplectic correctors. In G. Patrick, editor, *Integration Algorithms for Classical Mechanics.* Fields Institute, American Mathematical Society. To appear.

[11] M. Nelson, W. Humphrey, A. Gursoy, A. Dalke, L. Kalé, R. Skeel, K. Schulten, and R. Kufrin. MDScope—a visual computing environment for structural biology. *Comput. Phys. Commun.*, 91(1, 2 and 3):111–134, October 14 1995.

[12] C. S. Peskin and T. Schlick. Molecular dynamics by the backward Euler method. *Comm. Pure and Appl. Math.*, 42:1001–1031, 1989.

[13] A. Portillo and J. M. Sanz-Serna. Lack of dissipativity is not symplecticness. *BIT Numer. Math.*, 35(2):269–276, 1995.

[14] R. D. Skeel and J. J. Biesiadecki. Symplectic integration with variable stepsize. *Annals of Numer. Math.*, 1:191–198, 1994.

[15] G. Rowlands. A numerical algorithm for Hamiltonian systems. *J. Comput. Phys.*, 97:235–239, November 1991.

[16] J. Sanz-Serna and M. Calvo. *Numerical Hamiltonian Problems.* Chapman and Hall, London, 1994.

[17] T. Schlick and A. Fogelson. TNPACK—a truncated Newton minimization package for large-scale problems: I. Algorithm and usage. *ACM Trans. Math. Softw.*, 18(46):46–70, 1992.

[18] C. Scovel. On symplectic lattice maps. *Physics Letters A*, 159:396–400, August 1991.

[19] M. Tuckerman, B. J. Berne, and G. J. Martyna. Reversible multiple time scale molecular dynamics. *J. Chem. Phys*, 97(3):1990–2001, 1992.

[20] J. Wisdom, M. Holman, and J. Touma. Symplectic correctors. In G. Patrick, editor, *Integration Algorithms for Classical Mechanics*. Fields Institute, American Mathematical Society. To appear.

[21] G. Zhang and T. Schlick. LIN: A new algorithm to simulate the dynamics of biomolecules by combining implicit-integration and normal mode techniques. *J. Comput. Chem.*, 14:1212–1233, 1993.

DISCLAIMER

JUN 29 '93

93-7-1

Department of Energy
Washington, DC 20585

June 22, 1993

Professor Robert D. Skeel
2316 Digital Computer Laboratory
University of Illinois
1304 W. Springfield
Urbana, IL 61801

Dear Professor Skeel:

Your application entitled, "Numerical Methods for Molecular Dynamics," has been selected for funding under Applied Mathematical Sciences, subject to the completion of a grant. I recommend the authorization of a 18-month obligation of $75,000 in FY 1993 for the period of August 1, 1993 through January 31, 1995.

The negotiation of an appropriate grant is the responsibility of the Department's Chicago Operations Office. This letter is for information purposes only.

At the end of this funding period, the Department of Energy requires a final report describing the work accomplished under this grant, including a complete list of publications supported.

Should you wish to continue this work or other research with DOE support, a new application, subject to peer review, will be required at least six months before the proposed start date.

It is a pleasure to work with you in your research efforts.

Sincerely,

Fred

Frederick A. Howes, ER-30, GTN
Office of Scientific Computing
Office of Energy Research

cc: DOE Chicago Operations Office