

# Multiscale Research in Molecular Mechanics

*A. Brandt*

Department of Applied Mathematics and Computer Science  
The Weizmann Institute of Science  
Rehovot, 76100, Israel

## 1. Introduction

Molecular dynamics is a general and basic approach for performing atomistic-level simulations in order to understand, predict, design and control materials and processes. However, even with advanced supercomputers, current molecular dynamics algorithms severely restrict the modelling efforts to relatively small molecular systems and/or tiny evolution intervals, since their computing cost rises very steeply with problem size.

A sequence of preliminary model studies has indicated that the various types of scaling difficulties encountered in computational chemistry can in principle be overcome by multiscale algorithms. In collaboration with Dr. Ruth Pachter at the Materials Directorate, WL/MLPJ, Wright-Patterson AFB it has been decided to focus the present effort toward developing and demonstrating multiscale methods in molecular mechanics. Future, related efforts, depending on suitable support will include multiscale ab-initio calculations of electronic structures and the even more basic and accurate modeling of chemical reactions by multigrid Monte-Carlo methods for real-time Feynman path integrals.

This first report summarizes our long term objectives (Sec. 2) and our initial work on multiscale energy minimization methods (Sec. 3).

References [1]–[13] were used by us in getting acquainted with existing approaches in molecular simulations. We have also been delighted to discover several articles, such as [11]–[13], where some rudimentary forms of multiscaling have already been attempted, motivated by physical insights, with considerable gains. Unaware of several basic multiscale techniques, though, these attempts do not go far enough.

## 2. Long Term Objectives

The basic long term objectives of the project can be summarized as the following six tasks, ordered roughly according to the chronological order in which we have started or intend to continue to work on them, although much back and forth interaction between these coupled tasks is of course expected.

## 2.1 Fast summation of forces

Direct calculation of all the electrostatic interactions between  $N$  particles costs  $CN^2$  computer operations, where  $C$  is around 10. Instead, several methods exist to sum the forces in just  $C_1N$  operations (see, e.g., survey [GR]), although note that in three dimensions  $C_1 > 10^4$ , so these methods become advantageous only for  $N > 10^3$ . A multiscale method for fast evaluation, suggested in [OS] (based on an idea described earlier in [G2] and [TO]), will be used by us. It is based on a decomposition of the two-particle potential into a local part and a smooth part, the latter being evaluated at larger scales (interpolated from coarser grids), where a similar decomposition is being recursively used. One possible advantage of this approach is considerably smaller values of  $C_1$ . More important, the involved decompositions will give the kind of multiscale description of the force fields which is needed for the efficient multiscaling of the other tasks described below; see in particular the use of this decomposition in Secs. 3.2, 3.5 and 3.7 below.

Since we have already acquired enough preliminary experience with this force summation task, we will not develop it immediately further, until the need indeed arises in conjunction with our other tasks, described next.

## 2.2 Fast macromolecular energy minimization

This task serves two somewhat different objectives: one in statics, the other in dynamics. In statics, the objective is to calculate the lowest energy or the most stable conformations of large  $N$ -atom molecular structures; to find, in other words, the atom conformation  $r = (r_1, r_2, \dots, r_N)$  for which the potential energy  $E(r)$  is minimal. In dynamics, the objective is the solution of the system of equations arising at each time step of *implicit* dynamic simulations.

“Implicit” refers to the method which evaluates the forces (or the gradient of  $E(r)$ ) at each time step (partly or wholly) in terms of the particle *arrival* positions, i.e., positions at the *end* of the step. This method ensures stability of very large time steps, but it does not yield the arrival positions explicitly. Instead, they should be calculated by solving a large system of equations. (Also, this method damp molecular vibrations at scales not resolved by the large time step; we return to this point below.) Solving the implicit system of equations is equivalent to minimizing an *augmented* energy functional, identical to  $E(r)$  except for an additional quadratic *kinetic term* (cf., e.g., [6]). For large time steps this additional term is locally very small, but its large-scale effect is still profound.

Starting from a close enough first approximation to the desired minimal energy configuration, current minimization methods would converge to that configuration, in the case of a strongly coupled  $N$ -atom structure, in  $O(N^3)$  computer operations (not even including the operations for electrostatic force calculations, discussed

above). Multiscale methods promise to obtain a solution in just  $O(N)$  operations. This, however, requires a very careful development. Since this is our main *current* work, it is described in detail separately (Sec. 3 below). That description also includes a discussion of multiscale approaches to the harder situation where the given first approximation is *not* so close (Secs. 3.5–3.9).

### 2.3 Normal mode analysis

Near a given configuration of an  $N$ -atom molecular structure, the normal modes are the harmonic vibrations, i.e., the eigenfunctions of the harmonic (quadratic) approximation to the energy functional. They are important to approximate equilibrium and dynamic properties. In particular, the high-frequency modes may describe the vibrations not resolved by the large implicit time steps (see above), while low modes may describe important large-scale behavior.

The multiscale calculation of many low modes is expected to be extremely efficient, since it will be done on the *coarser levels* of the multiscale fast-solver described below (Sec. 3). Moreover, on those coarser levels one can also directly calculate the combined action of all these modes, and thus, due to the FAS version (see Sec. 3.6), calculate large-scale *anharmonic* effects as well. (This becomes rather similar to performing many large-scale statistical passes: see item (d) in Sec. 2.4.)

We plan to demonstrate such basic capabilities after the multiscale solver (discussed in Sec. 3 below) is sufficiently developed.

In a less obvious way, a multiscale structure can also be very efficient in calculating higher modes, including the highest ones, since the separation between increasingly closer modes can be done on increasingly coarser levels. Simple models indicate that a well structured multiscale eigenbase for an  $N$ -atom molecule may be calculated in as little as  $O(N \log N)$  operations. Such a calculation is, however, rather complicated, and its importance is questionable, because the exact large-scale separation of close *high-frequency* modes may have little physical significance. Probably more relevant to the high-frequency behavior are the statistical approaches, discussed next (Secs. 2.4 and 2.5).

### 2.4 Monte-Carlo methods at equilibrium

In equilibrium at temperature  $T$ , the probability of each configuration  $r$  to appear is proportional to  $P(r) = \exp(-E(r)/T)$ . To calculate equilibrium statistics, an atom-by-atom Monte-Carlo process is usually performed. In this process, each atom in its turn changes position stochastically, according to the probability density distribution  $P(r)$ . Making repeated sweeps of this process, one can

calculate the desired statistics on the sequence of produced configurations.

To calculate accurate averages of some observable, however, an extremely long sequence of configurations is needed. There are two basic reasons for this complexity: (1) Due to the local nature of the Monte-Carlo process, only very slowly it affects large-scale conformational features, hence extremely many Monte-Carlo sweeps are needed to produce each new, statistically independent configuration. (2) Many such independent samples are needed to average out the deviation observed at each of them.

For some model problems, multigrid Monte-Carlo algorithms were developed which overcome *both* these complexity reasons (see [BG] and [L1]). The algorithms are similar to the multiscale cycles described below (Sec. 3), with the following four modifications.

(a) The energy functional should be the true potential  $E(r)$ , not the augmented one (cf. Sec. 2.2).

(b) The Gauss-Seidel relaxation (atom-by-atom minimization) sweeps should be replaced by atom-by-atom Monte-Carlo sweeps.

(c) The approximation of the Hamiltonian (energy functional)  $E$  in the coarsening process should be done in a stochastic manner, to retain the statistical fidelity (the “detailed balance”). Methods to achieve that are highly nontrivial, and may require careful research and development. However, in view of the approximate nature of the molecular-mechanics Hamiltonian to begin with, exact detailed balance may not be required, as long as statistical fidelity is retained in the limit of very smooth fluctuations. This can be achieved much more easily.

(d) The multiscale cycle should switch many times back and forth between coarse levels, before returning to finer levels. In this way many samples of large-scale features can be averaged over. Not so many passes are needed at the finer scales, because many fine-scale features are already present, and hence averaged over, in any one configuration.

## 2.5 Small-scale statistics with large-scale dynamics

The multiscale structure allows the combination of statistical simulations at small scales with time-accurate dynamics at large scales. For this purpose the multiscale solver (Sec. 3) should be modified in two ways.

First, the time-step discretization should be such that it gives accurate (non-damping, energy conserving) approximations for all scales whose time-accurate dynamics need be simulated.

Secondly, at all finer scales (finer levels of the multiscale solver), the Gauss-Seidel relaxation sweeps should be replaced with Monte-Carlo sweeps, as described above (Sec. 2.4).

This scheme comes close to the one described, with a slightly different motivation, in Sec. 3.9 below.

## 2.6 Material “homogenization”

The coarse-level energy functional of a well-developed multiscale solvers should effectively yield the large-scale behavior of the simulated material. This would yield the crucial link between the atomistic level description and the continuum-level material modeling, thus enabling a basic understanding and design of phenomena at various scales.

Indeed, to simulate a material at macroscopic scales, a multiscale processing needs not resolve the atomic scales over a macroscopic domain. Rather, a tiny domain (but still very large compared with the smallest inter-atomic distances) can first be simulated. This simulation should employ 4D (space + time) periodic boundary conditions, using a 4D multiscaling; i.e., periodicity and coarsening is used not only in space, but also in time. (The optimal adjustment of the period size in each space direction and in time can very inexpensively be performed at the coarsest levels of such a solver.) In the next stage the periodicity is used to extend the domain (e.g., double the period size in each direction), while the description is *coarsened* (i.e., the finest scale of the multiscale solver is dropped, retaining the energy functional it supplied to the coarser levels). Then a simulation with this coarsened descriptions erase the original periodicity, retaining periodicity only at the extended domain boundaries. Repeating this procedure, the algorithm can simulate ever larger domains with ever coarser descriptions, until macroscopic scales are attained.

We hope to demonstrate simple instances of such multiscale expansions in some, much later stages of the project.

## 3. Energy Minimization

The first crucial step in constructing a fast multiscale energy minimizer is to design its inner cycle: such that it yields fast convergence near the minimum, i.e., for sufficiently close first approximations. We discuss this cycle first (Secs. 3.1–3.4). Later (Secs. 3.5–3.9) we will discuss the nonlinear aspects of driving the cycles when the initial approximation is not so good.

When linearized, the molecular energy minimization problem is somewhat similar to the minimization problem encountered in structural mechanics, for which very efficient multigrid solvers have been developed. Of those, the closest to the ones needed in molecular mechanics are the *algebraic multigrid* (AMG) solvers [A1]–[A4], which do not assume that the problem arises from a continuum structure

or that the unknowns are really placed on a grid. Some of the basic ideas of the AMG solvers, as well as the experience obtained with them, will be important in our present development.

### 3.1 Main difficulty and a model

There are several types of computational difficulties peculiar to molecular energy functionals. To develop an efficient multiscale solvers we must start by studying each of these difficulties in separation from the others, by treating appropriate model problems.

The main special difficulty arising in (linearized) molecular structures is that different kinds of forces are associated with very different coupling strengths (manifested for example in different time scales of the corresponding dynamics). This has profound implications for the multiscale cycle.

To explain and develop the principles of efficient cycles under such situations, we have started with a simple 2D model problem: A chain of  $N$  atoms at the planar positions  $r_i = (x_i, y_i)$ , ( $i = 1, \dots, N$ ), with the energy functional

$$E(r) = \sum_{i=2}^N B_i (|r_i - r_{i-1}| - b_i)^2 + \sum_{i=2}^{N-1} K_i (\theta_i - \theta_{i0})^2,$$

where  $\theta_i$  is the angle between the vector  $r_{i+1} - r_i$  and the vector  $r_i - r_{i-1}$ , and where the bond constants  $B_i b_i^2$  are much larger than the angle constants  $K_i$ . To make the problem nontrivial, some of the atom positions (e.g.,  $r_1$  and  $r_N$ ) may be fixed. This two-dimensional model mimics the three-dimensional situation where dihedral angle couplings are much weaker than those of both bonds and bond angles. (We will of course insist on applying to this model methods which do not depend on its special, non-representative features. For example, we will avoid using the bond lengths and angles as the dependent variables, which would yield a straightforward separation between weak and strong interactions, but would not be applicable in more general situations. Similarly we will refrain from solution techniques tailored only for one dimensional problems. For example, the technique of [11] can be very effective for such problems, but much less generally).

In the presence of such highly nonuniform couplings, the simple atom-by-atom minimization procedure (equivalent to Gauss-Seidel relaxation) is extremely inefficient. For example, in the model problem, if the  $K_i$ 's are comparable to  $\varepsilon$  times the typical size of  $B_i b_i^2$ , with some  $0 < \varepsilon \ll 1$ , then the convergence of an atom-by-atom relaxation process would require  $O(N^3/\varepsilon)$  computer operations. Since coupling strengths in molecules range over at least four orders of magnitude,  $\varepsilon \approx 10^{-4}$  can be considered typical.

For small  $\varepsilon$ , the naive multigrid solvers (using the usual coarse-to-fine interpolation schemes) would also be quite ineffective, requiring  $O(N/\varepsilon)$  operations. By

contrast, we believe that the general multigrid principles outlined below should solve any (harmonic) molecular energy minimization in just  $O(N)$  operations, independently of coupling strength ratios. We describe below these general principles and, as an example, their specific application to the model problem.

### 3.2 Relaxation

The relaxation is a local process whose purpose in multiscale solvers is just the fast reduction of “non-smooth” errors, or their fast replacement by “smooth” ones. A given error function is considered “non-smooth” (or “locally reducible”) if it is associated with relatively large residual forces, where “relatively large” means that they are comparable to the largest residual forces producible by any error comparable in magnitude to the given one.

The usual Gauss-Seidel (GS) relaxation (or atom-by-atom minimization) is actually very efficient in reducing non-smooth errors.

The only concern is the amount of work in calculating long-range (e.g., electrostatic forces). In the GS relaxation, the forces summed for each atom movement are based on the most recent locations of all other atoms. For long range forces, such a separate summation for each atom would be too expensive. On the other hand, the Jacobi-type relaxation, where the forces are based on atom positions at the beginning of the relaxation sweep and can therefore be based on efficient simultaneous summation of all forces, has smoothing properties much less satisfactory and requires carefully chosen under-relaxation parameters.

The best scheme seems indeed to be a combination of the Gauss-Seidel and Jacobi schemes, based on the decomposition of forces into a local part and a smooth part (see Sec. 2.1 above): The smooth part of the forces can be updated once per sweep, while the local forces should be updated during relaxation. Actually, since the local forces greatly diminish toward the margin of their support, it is suspected that only a fraction of them, those closest to the affected atom, need to be updated inside the relaxation sweep.

An interesting further possibility we want to explore is to entirely neglect the smooth part of the forces during relaxation, since its significant effect is only on smooth errors. This would mean that the smooth part of the forces need not even be interpolated to the atom positions, implying a substantial work saving.

### 3.3 Coarse-level correction

The errors which are not reduced efficiently by the atom-by-atom minimization sweeps must exhibit relatively small residuals, as defined above. Hence such errors must be “smooth”, or, more precisely, they nearly belong to the subspace

spanned by the lower eigenfunctions (of the linearized system). They can thus be described by a smaller number of degrees of freedom. These degrees of freedom form the (first) “coarse level”.

In the case of multigrid solvers for isotropic elliptic PDEs discretized on a uniform grid, relaxed errors (i.e., errors left after a couple of Gauss-Seidel relaxation sweeps) must be smooth in the usual sense (a grid function having small local differences compared with the function itself). Hence such errors can well be approximated by functions interpolated from a *coarser grid*. Thus the degrees of freedom of the “coarse level” are those of a coarser-grid function. This function is calculated by solving the “coarse-grid equations”, and its interpolation is then used as a correction to the former solution on the original (fine) grid. The coarse-grid equations themselves are determined by the requirement that the resulting interpolated correction lowers the energy as far as possible.

In the case of molecular structures, the sense of “smoothness” of relaxed errors is more complicated, due to the unevenness of the couplings and the lack of a uniform grid structure. Hence both the coarse level variables and the interpolation from them to form the corrections to the atom positions are also more complicated.

*The coarse-level variables* always represent displacements, i.e., changes, in atomic position (since they are designed to approximate the errors). One approach for choosing these variables is, as in AMG, to simply choose a *subset* of all the displacements of the original variables. We tend to take this approach for the first coarsening step discussed here. (By contrast, at coarser levels, where movements will no longer be constrained by very uneven couplings, the next-coarser level variables should probably best be defined on a 3D spatial *grid*, especially in the (usual) case where the molecules, including solvent, are not one dimensional, but fold and fill the full three dimensional space.)

*The interpolation* should directly reflect the strongest couplings. For example, in the above 2D model problem, the coarse level variables may simply include the displacement (vector) of every other atom, say those with even indices. The interpolation to the odd-numbered atoms is uniquely determined by the requirement that, to a first order, the bond lengths are unchanged.

In a more complex situation, the most general approach is to find the “sense of smoothness” of the relaxed errors by a local computation of the lowest eigenmodes. By this we mean, for example, computing the lowest eigenmodes of a system which only includes the interactions between the atoms inside a certain local box. In some interior subdomain of that box, a good coarse-to-fine interpolation can then be determined by the requirement that it is fully compatible with (exactly satisfied by) each of the lowest eigenmodes.

Although local in principle, this general procedure for deriving the interpolation is still quite expensive. Fortunately, at the finest (atomistic) level, a more explicit choice, based (as in the example) on the strongest couplings, will usually be quite straightforward and effective. (When the coarsening process is later ap-



plied at a coarser base level, with more complex interactions, it may be important to employ the general procedure, but its cost then will be much lower, due to the smaller number of variables at that coarser base level.)

Once the coarse-to-fine interpolation has been chosen, the *coarse-level energy functional* follows directly from the given (fine-level) energy functional. More precisely, a polynomial approximation (a Taylor expansion) for the latter (around the current fine grid solution) yields a polynomial (of the same degree) approximation for the former, since interpolation is a linear operator.

After minimizing this coarse-level energy functional (by the method discussed in the next section), the minimizing function is interpolated to the fine-level and used as a correction to the previous solution there. Then the process may be repeated: some Gauss-Seidel relaxation sweeps are made, followed by a new calculation of the coarse-level energy functional, its minimization, and interpolation of the minimum to further correct the atomic positions. And so on, producing fast converging iterations.

### 3.4 Recursion: multilevel cycle

The method for fast approximate minimization of the *coarse-level* energy is the same as that we have described for the original (atomistic) level. Namely, a couple of Gauss-Seidel relaxation sweeps (now on the coarse level) are followed by calculation of an energy functional for the next coarser level, (approximate) minimization of the latter, interpolation of the obtained minimum as a correction to the (first) coarse-level configuration, and perhaps some additional Gauss-Seidel sweeps for the latter. In this way we can recursively define the *multilevel* (or “*multiscale*”) *cycle*.

The main remaining task in constructing the cycle is to choose, at each *current* level, the next-coarser-level variables and the coarser-to-current interpolation operator. These choices will determine the efficiency of the algorithm. Some of the main approaches for making them were indicated above. The development and testing of versatile and efficient coarsening schemes will be our major task during at least the first research year.

### 3.5 Nonlinear aspects

So far we have dealt with fast convergence to the minimum of polynomial (e.g., quadratic) approximations to the energy functional. To be sure, the atom-by-atom relaxation process is actually performed with the *full* energy functional (local-Newton Gauss-Seidel relaxation); but the coarsening process is done in terms of a polynomial (Taylor) expansion of the energy functional around some current

configuration. The expansion does not really assume small displacements: In terms of *local* forces (including the local part of the electrostatic force, in the sense of the decomposition mentioned in Sec. 2.1), the expansion only assumes small *strains*, i.e., small displacement *differences*, or, in other words, *smooth* displacements. Since what we want the coarse level to perform are just such smooth moves, the expansion holds true even for large displacements (except for the smooth part of the electrostatic forces — discussed below).

However, when the initial configuration is not close enough to the desired minimum of the full energy functional, just making a couple of iterations of polynomial approximations (e.g., a couple of global Newton steps) may not work. We plan several kinds of multiscale devices to deal with the nonlinear, more global convergence; they are described in the next sections (Secs. 3.6–3.9).

### 3.6 FAS cycles

One device is the FAS (“Full Approximation Scheme”) version of multigrid algorithms. In this scheme, the coarse-level variables are shifted: instead of using the *displacement* variables (those which, upon interpolation, describe atom displacements), each FAS variable is the *sum* of a displacement variable and the corresponding atom location just before the coarsening process. Since that location has been assumed (in the polynomial energy expansion) to be fixed throughout the solution processes for the coarse level (including the processes on still coarser levels), this shift is just a simple additive constant shift for each variable. (Note that the FAS variable actually describes the full new location of the atom; hence the term “full approximation” scheme. But the interpolation to atoms not represented on the coarse grid is still done in terms of the *displacements*, which are calculated by back-shifting the FAS variables.) The advantage of the FAS variables is that they allow to write the coarse-level equations in a special way, such that the approximation holds for a much wider range of coarse-level displacements (see detailed description, for the case of PDEs, in Chapter 8 of [G2]). In essence, the FAS allows taking into account nonlinear (anharmonic) interactions at the coarse levels, since it retains knowledge of the full approximation, not just its displacements.

This in particular applies to the smooth part of the electrostatic forces: In term of the FAS variables that part can retain its form in the coarse level. Then later, on that coarse level, this smooth part will again be decomposed into a local part and a smooth part, “local” and “smooth” now in a larger-scale sense; and in the next coarsening, from the first coarse level to the next coarser one, it will be that *new* (larger-scale) smooth part that will retain its form. And so on, to increasingly coarser levels with increasingly smoother parts of the electrostatic interaction. This will make the energy Taylor expansions (cf. Sec. 3.5). valid for large (smooth) displacements even in terms of the electrostatic forces. (This also

reinforces the importance of the smooth + local decomposition of electrostatic forces, as discussed in Sec. 2.1 above.)

### 3.7 FAS-FMG algorithm

The actual multiscale solver that will be used at each time step of molecular dynamics is the *FAS-FMG solver* (or the “*F* cycle”; described, e.g., in [PR]). In it, an (approximate) energy functional is first transferred to increasingly coarser levels, in terms of FAS variables. Then a so-called FMG (= “full multigrid”) solver is applied. This solver starts at the coarsest level and sequentially proceeds to increasingly finer “base levels”. At each base level, a first approximation is obtained by interpolating the displacements from the previous (next coarser) base level. Then the approximation is improved by several multilevel cycles, of the type described in Sec. 3.4, except that they start at the current base level, not at the original (finest, atomistic) level. Then the solution is interpolated to the next, finer base level; and so on, until the finest level (of atoms) is reached and multilevel cycles are performed for it.

Actually, the experience with PDEs is that only *one* multilevel cycle at each base level is usually enough, since the initial approximation, obtained from the next coarser level, is already very good. One multilevel cycle is usually also enough at the finest level, for the same reason, making this solver very inexpensive. Effectively, the FAS-FMG algorithm makes global-Newton iterations unnecessary.

### 3.8 Continuations

A general powerful technique to deal with difficult nonlinear problems is to combine the FMG process with a *continuation (parameter embedding)* process (see Sec. 8.3.2 in [G2]). This means that at the initial stages of the FMG solver (at multilevel cycles for coarse base levels) some problem parameters are changed, to allow large scale movements not to be excessively affected by small scale details. As the FMG process advances to even finer base levels, the problem parameters are gradually restored, to finally reach their true value when the algorithm reaches the finest (the target, atomistic) level.

In particular, a natural FMG-continuation process for molecular mechanics is to ignore, at each base level, the local part of the electrostatic interactions of the next finer base level. That is, in the formulation of the base level energy functionals *in the beginning of the FAS-FMG algorithm*, this part is omitted; it is however *not* omitted in any coarsening within any of the multilevel *cycles*.

A related technique is to use a continuation process with a sequence of *several* multilevel cycles for the *same* base level, the finest base level in particular. Natural

continuation parameters in this respect are the stiffness coefficients of the strongest couplings. For example, in the above 2D model problem (Sec. 3.1), the values of  $B_i$  may be softened (lowered) when the approximate energy functional of the coarse level is calculated (see end of Sec. 3.3; this softening is especially important in case higher-than-quadratic Taylor expansions are used). As convergence is approached at subsequent cycles, the softening can be gradually taken out. The full-strength (unsoftened) coefficients should still be used in all the *relaxation* sweeps between coarsening steps. The softening at each coarsening step is such that it allows the expected size of the large-scale displacements not to be stifled by the strong couplings. (Although the effect of the latter vanishes to a first order if interpolation is designed as described in Sec. 3.3, the second-order effect may still be stifling if the coupling strength ratio is very large.) Since such a stifling (being a second order effect) disappears for sufficiently small displacements, the softening can indeed be gradually taken out as convergence is approached.

### 3.9 Stochasticity, combined with dynamics

A powerful method to escape local energy minima in search of a more global minimum is to add stochasticity to the minimization process. *Simulated annealing* is the most well known example. Since in molecular mechanics problems a multitude of local minima comes with multiscale attraction basins, a multiscale version, called *multilevel annealing* (see Sec. 3 of [RA]), can be much more effective (as shown in [RA] for spin glass problems).

The energy functional discussed thus far is the *augmented* one, including the quadratic *kinetic term* (see Sec. 2.2 above). This augmented energy would be used in the annealing processes at each time step. For the dynamics of large biomolecules, however, the most natural stochasticity, and possibly the only one to guarantee correct dynamics, is the one associated with the true temperature of the material. This suggests a multiscale process where on finer scales (those scales associated with vibrations not resolved by the time step), instead of the Gauss-Seidel relaxation, a Monte-Carlo process (cf. Sec. 2.4) will be applied. It will use the relevant physical temperature and the bare (not augmented) potential energy. This process comes in fact close to that of Sec. 2.5, except that it may be simpler because it can gloss over the problem of detailed balance. Also, it may be desired to finish the process off by an annealing phase (letting the temperature decrease gradually to zero).

## References

- [1] L. Verlet, Computer ‘experiments’ on classical fluids, I. Thermodynamical properties of Lennard-Jones molecules, *Phys. Rev.* **98** (1967).

- [2] A. Rahman, Correlations in the motion of atoms in liquid Argon, *Phys. Rev.* **136** (1964) A405.
- [3] M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, New York (1990).
- [4] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan and M. Karplus, CHARMM: A program for macromolecular energy, minimization and dynamics calculations, *J. Comput. Chem.* **4** (1982) 187.
- [5] W.F. van Gunsteren, H.J.C. Berendsen and J.A.C. Rullmann, Stochastic dynamics for molecules with constraints Brownian dynamics of *n*-alkanes, *Molecular Phys.* **44** (1981) 69.
- [6] C.S. Peskin and T. Schlick, Molecular dynamics by backward-Euler method, *Comm. Pure Appl. Math.* **42** (1989) 1001.
- [7] A.M. Nyberg and T. Schlick, Increasing the time step in molecular dynamics, *Chem. Phys. Lett.* **198** (1992) 538.
- [8] P. Derreumaux, G. Zhang, T. Schlick and B. Brooks, A truncated Newton minimizer adopted for CHARMM and biomolecular applications, *J. Comput. Chem.* **15** (1994) 532.
- [9] T. Schlick, S. Figueroa and M. Mezei, A molecular dynamics simulation of water droplet by the implicit Euler/Langevin scheme, *J. Chem. Phys.* **94** (1991) 2118.
- [10] G. Zhang and T. Schlick, The Langevin/implicit-Euler/normal-mode scheme for molecular dynamics of large time steps, *J. Chem. Phys.* **101** (1994) 4995.
- [11] M.H. Hao, S.C. Harvey, Analyzing the normal mode dynamics of macromolecules by the component synthesis method, *Biopolymers* **32** (1992), 1393.
- [12] M.E. Tuckerman, B.J. Berne and A. Rossi, Molecular dynamics algorithm for multiple time scales: systems with disparate masses, *J. Chem. Phys.* **94** (1991) 1465.
- [13] M.E. Tuckerman and B.J. Berne, Molecular dynamics in systems with multiple time scales: systems with stiff and soft degrees of freedom and with short and long range forces, *J. Chem. Phys.* **95** (1991) 8362.
- [A1] A. Brandt, S. McCormick and J. Ruge, Algebraic multigrid (AMG) for automatic multigrid solution with application to geodetic computations, Institute for Computational Studies, POB 1852, Fort Collins, Colorado, 1982.
- [A2] A. Brandt, S. McCormick and J. Ruge, Algebraic multigrid (AMG) for sparse matrix equations, in: *Sparsity and its Applications* (D.J. Evans, ed.), Cambridge University Press, Cambridge, 1984, pp. 257-284.

- [A3] A. Brandt, Algebraic multigrid theory: The symmetric case, *Appl. Math. Comp.* **19** (1986) 23–56.
- [A4] J. Ruge and K. Stüben, Algebraic multigrid, in: *Multigrid Methods* (S.F. McCormick, ed.), SIAM, Philadelphia, 1987, pp. 73–130.
- [BG] A. Brandt, M. Galun and D. Ron, Optimal multigrid algorithms for calculating thermodynamic limits, *J. Stat. Phys.* **74** (1994) 313–348.
- [G2] A. Brandt, Guide to multigrid development, in: *Multigrid Methods* (W. Hackbusch and U. Trottenberg, eds.), Springer-Verlag, 1982, pp. 220–312.
- [GR] L. Greengard, Fast algorithms for classical physics, *Science* **265** (1994) 909–914.
- [L1] A. Brandt, Multigrid methods in lattice field computations, *Nucl. Phys. B* (Proc. Suppl.) **26** (1992) 137–180.
- [OS] A. Brandt, Multilevel computations of integral transforms and particle interactions with oscillatory kernels, *Comp. Phys. Comm.* **65** (1991) 24–38.
- [PR] A. Brandt and J. Greenwald, Parabolic multigrid revisited, in: *Multigrid Methods III* (W. Hackbusch and U. Trottenberg, eds.), Birkhäuser Verlag, Basel, 1991, pp. 143–154.
- [RA] A. Brandt, D. Ron and D.J. Amit, Multi-level approaches to discrete-state and stochastic problems, in: *Multigrid Methods II* (W. Hackbusch and U. Trottenberg, eds.), Springer Verlag, 1986, pp. 66–99.
- [TO] A. Brandt and A.A. Lubrecht, Multilevel matrix multiplication and fast solution of integral equations, *J. Comp. Phys.* **90** (1990) 348–370.