

A FUTURE ROLE FOR NUMERICAL AND APPLIED MATHEMATICS IN MATERIAL SCIENCES*

S. ABARBANEL[†], S. TSYNKOV^{†‡}, AND E. TURKEL[†]

Abstract. In this report, we give a general review of the possible areas, in which the methods of applied mathematics may be implemented to the benefit of modern material sciences. In particular, we address the emerging framework of nano-technologies, and discuss both the issue of modeling, as well as that of solving (typically, by approximate/numerical methods) the mathematical problem as presented by the model. We also emphasize the crucial role of close collaboration between the mathematicians on one side, and scientists and engineers on the other side, for the overall success.

Key words. multiple scales, stiffness, ergodicity, molecular dynamics, rapid evaluation of forces, long-range and short-range forces, fast multipole expansions, field potential, far-field boundary conditions, multi-grid methods, truncation criteria, mathematics-based and physics-based modeling

Subject classification. Applied and Numerical Mathematics

1. Introduction. We present this document to identify areas in which modern material sciences may benefit from the application of the methods in numerical and applied mathematics. Some examples include the branches studying nano-structures and those deriving constitutive equations for novel materials. We fully recognize the centrality of powerful modern computers, and this underlines much of what we present. However, the role of numerical analysis, and other analytical mathematical methods, cannot be underestimated.

This document is written by applied mathematicians and so reflects the way that mathematicians approach problems. We do not claim sufficient expertise in the physical and chemical foundations of material science, rather we try to address issues related to the mathematics underlying the existing models. Hence, we propose to improve the computational efficacy of the corresponding solution methodologies. We hope that in certain instances future mathematical analysis will lead to the modification of the models, and to the development of alternative computational procedures.

As in the previous well-known cases, such as computational fluid dynamics (CFD), computational acoustics, and computational electromagnetics (CEM), mathematicians usually follow in the footsteps of scientists and engineers that are more familiar with the physical applications. On the other hand, the rigorous way of looking at things, which characterizes mathematics, often helps scientists and engineers to focus on key issues that, once resolved, lead to major research breakthroughs. On some occasions, that cannot be predicted, advances in mathematics truly revolutionize the corresponding applied field. This has happened, for example, in CFD, an area that was changed profoundly following mathematical advances in the theory of conservation laws. Interactions between mathematicians and scientists in a particular applied field may also lead to adopting some general mathematical techniques that have been well established and successful in other areas. This has happened, for example, in CEM, with the introduction of the time-domain finite-difference methods in this field, which supplemented the previously used frequency-domain techniques.

*This research was supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-97046 while the authors were in residence at ICASE, NASA Langley Research Center, Hampton, VA 23681-2199, USA.

[†]Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel.

[‡]North Carolina State University, Box 8205, Raleigh, NC 27695, USA.

The present document has its roots in informal interactions between applied mathematicians and material scientists and engineers at ICASE and NASA LaRC. Those interactions, which are still ongoing, include seminars as well as less formal discussions. It is our hope therefore that this paper will lead to a blueprint of a multi-year joint research effort at ICASE and NASA Langley Research Center. This effort should include residence at ICASE of a group of researchers composed of both numerical/applied mathematicians and theoretical material scientists working in close collaboration. We clearly realize the importance of incorporating input from, and providing feedback to, the experimentalists at NASA Langley, and possibly other laboratories.

Even though at the current stage it is difficult to predict what areas of material science may benefit most from such a collaborative effort, we will try to delineate those areas that have two characteristics:

- They are, in our opinion, promising from the scientific point of view.
- They are particularly suitable and appropriate for the NASA LaRC material science environment.

2. Resolving the issues related to multiple temporal and spatial scales. It is known that the existence of multiple, often highly disparate, time scales is a typical situation for complex multi-atom and multi-molecule models. Different scales usually originate from different mechanisms of interaction between the particles, and are determined by different types of strain put on the bonds between the particles within a given ensemble. Stretching typically results in the fastest oscillation, which is followed by bend, torsion, and finally motion of non-bonded particles. The performance of a standard explicit numerical integration routine when applied to the system of ordinary differential equations (ODEs) that represents the Newton's second law and governs the motion of the particles, will be limited by the rate of fastest oscillations. This leads to very small time steps to avoid numerical instabilities. Small time steps, in turn, imply expensive computational effort and severely limit the overall model time for which the system may be integrated even on the most advanced modern computer systems during a reasonable (days, weeks, more rarely months) wallclock time. In typical molecular dynamics (MD) simulations, the aforementioned model time is on the order of nanoseconds, whereas substantially longer time intervals are often required to make practical predictions.

This problem has been known for a while and recognized in the literature. In mathematical terms, the property of having highly disparate time scales, with the fastest one determining the performance of the solver, is known as stiffness. It can be conveniently characterized by the condition number of the matrix of the corresponding system of ODEs. For stiff systems the condition number, which is the ratio of the largest to the smallest eigenvalue, is large. Eigenvalues characterize the rates, or speeds, of the corresponding processes.

Most of the approaches that have been tried, to date, to alleviate the effect of stiffness on the computational performance in MD simulations for nano-materials have a physical origin. These approaches are equivalent to analyzing different physical mechanisms and the corresponding time scales. This leads to a subsequent decision to use an approximate, rather than a completely accurate, representation of a particular mechanism or several mechanisms that are most hampering the performance. For example, the fastest oscillation can be "frozen". Some of these approaches can be reformulated in mathematical terms, in which case they reduce to techniques developed in the framework of the numerical analysis of ODEs, which has its own extensive history of dealing with the issue of stiffness. However, many other methods for stiff systems available in numerical analysis of ODEs do not relate directly to the physical origins of the problem, but are built primarily to take into account the expected properties of the solution rather than its driving mechanisms.

In our view, it will be extremely beneficial if the problem of stiffness in MD simulations is addressed by a multidisciplinary team of physicists/chemists/material scientists on one hand, and numerical analysts/applied mathematicians on the other hand. We hope that this collaboration will fuse the hints provided by the physics with the mathematical imperatives to create optimal approaches to the issues of stiffness in MD simulations.

In the spatial domain, the nano-structure is typically a part of a much larger media. Besides the properties of the nano-structure one desires to know its impact on the larger structure and conversely, the impact of the media (surrounding matrix) on the nano-structures. These structures do not behave the same way in isolation or when embedded in a larger media. Furthermore, the nano-structures are frequently combined to form larger structures. However, any analysis of such interactions requires one to investigate many orders of magnitude in scale between the nano-structure and the macro-structure in which it appears. In the last few years several mathematical tools based on multigrid, multi-resolution, and wavelets have appeared that allow one to combine phenomena on different space scales. In addition, let us note that the issue of disparate spatial scales may also arise in the context of far-field boundary conditions, see Section 5.

3. Dealing with the problem of non-ergodicity. Many systems of molecules, such as glassy materials, meta-stable states, and nearly harmonic solids, are not ergodic. Since the computation of time average quantities is a lot less expensive than the evaluation of ensemble averages, most MD calculations are performed on a single system and extract the required average quantity (say, density) by performing temporal averaging. The ergodic hypothesis is then invoked to justify the result as being the correct one. The fact that many systems are in fact non-ergodic means that the time-average quantities computed via the usual MD simulations are incorrect. It may well behoove us to study carefully, and quantitatively, the effects of non-ergodicity.

Is it possible to define quantitatively (at least for a given system) the extent of deviation from ergodicity? Can this deviation be correlated, again quantitatively, with the error resulting from using time-averaging?

We feel that this issue would be a worth-while topic of basic research that can be tackled best by a cross-discipline group as envisioned in this document.

4. Dealing with long-range forces in molecular dynamics. One can describe molecular dynamics simulations as a way to numerically integrate a large collection of ordinary differential equations (ODEs) that represent the law of motion (Newton's second law) for all the particles that comprise the system under study. The source terms for these ODEs are forces acting on the particles. These forces typically originate from interactions between the particles. For a system of N particles such that every particle interacts with all others, a direct evaluation of forces will require computing $\mathcal{O}(N^2)$ interactions that are determined by the properties of individual particles (e.g., electric charges) and by dynamically changing distances between them. Clearly, the required computational effort per force evaluation will scale quadratically with the number of particles in the system. And because the evaluation of forces is needed on every time step of the numerical ODE solver, the corresponding computational algorithm will quickly become unacceptably expensive with the increase of N .

Depending on the type of interaction (i.e., its physical nature), the forces acting between particles can be classified as either long-range or short-range. For the former, any meaningfully defined characteristic length of interaction will be comparable to the size of the domain of interest, whereas for the latter it will be much smaller than the domain size. We postpone the discussion on short-range forces till Section 7. Here we focus on the long-range forces, which are typically of Coulomb nature in the context of molecular dynamics. From the standpoint of numerical simulation of the motion of the particles, the key property of these forces is

that they cannot be disregarded even at large distances and therefore, the $\mathcal{O}(N^2)$ computational complexity associated with the direct evaluation of long-range forces on every time step becomes a major bottleneck of the numerical algorithm.

In recent years, an efficient algorithm has been proposed for calculating forces between interacting particles at an $\mathcal{O}(N \log N)$ effort rather than $\mathcal{O}(N^2)$ expense. These algorithm employs special hierarchical summation rules based on fast multipole expansions for Coulomb-type potentials. This algorithm, however, has not yet gained a solid ground in molecular dynamics simulations, especially as they apply to studying the nano-structures. On the other hand, we expect that full-fledged implementations of fast multipole summation rules in molecular dynamics codes, even though it will undoubtedly be a rather involved task, may provide a true breakthrough from the standpoint of reducing the execution times; and as such we believe it is certainly worth the effort. Moreover, no matter how far away from one another the following two groups of methods may seem to be, the mathematics behind the fast multipole algorithms is, in fact, very close to another class of efficient numerical techniques known as multigrid. Multigrid methods provide $\mathcal{O}(N \log N)$ computational procedures for solving partial differential equations on the grid. It is our belief that the similarities between the two groups of methods are worth exploring thoroughly, since in the context of molecular dynamics these methods may efficiently complement one another. Indeed, an alternative, and sometimes preferable, way to evaluate the forces acting on the particles is to first solve the differential equation for the field potential, which clearly calls for the application of a multigrid technique.

The ideas resembling one specific building block of fast multipole schemes have already been adopted in many molecular dynamics algorithms, although apparently without direct relation to, and completely outside of, the fast summation rules content. We refer to the idea of clustering, i.e. treating groups of particles as one “super-particle” for the purpose of evaluating forces, etc., far away from it. As has already been mentioned, fast multipole methods use hierarchical systems of such clusters, which allows them to use efficient $\mathcal{O}(N \log N)$ summation rules for the evaluation of the forces. The partitioning of particles into groups, sub-groups, etc., in fast multipole methods is done formally, based on pure mathematical reasoning. In contradistinction to this, the clustering that is currently in use in molecular dynamics simulations is based primarily on physical arguments. However, it does not create hierarchical systems of clusters and so does not obtain the full benefit of the fast summation. As such, we believe that the connections between the two approaches to clustering are worth careful exploration. This has the potential payoff of being able to obtain on one hand more physics-friendly fast summation schemes and on the other hand, simple improvements to the existing clustering techniques in molecular dynamics codes by using more mathematical insight.

5. Studying the role of boundary conditions. Numerical simulations in material sciences that involve nano-scales typically employ periodic boundary conditions at the outer boundaries. Perhaps the most common computational setup that is introduced primarily for reasons of convenience include a particular nano-structure or interest surrounded by a region filled with polymer molecules (the so-called matrix). To obtain a finite-dimensional computer model, the system of equations solved inside this region should be closed, which is done using the periodic boundary conditions. In the molecular dynamics perspective, these boundary conditions mean that the particles that leave the domain on one of its sides immediately re-enter it on the opposite side. A key advantage of this approach is its simplicity and self-sufficiency. Often, this is the only straightforward way to set the external boundary conditions. However, besides limiting the admissible domain shapes to parallelepipeds and their equivalents of some kind (otherwise, periodic boundary conditions

cannot be imposed¹), setting the periodic boundary conditions at the outer boundary often raises concerns of a more fundamental nature.

The foremost concern is the influence of the treatment of the outer boundary on the results of the simulation inside the computational domain. Adopting the conventional terminology of “near field” (the nano-structure of interest and the matrix in its immediate vicinity) and “far field” (the matrix away from the nano-structure), we can formulate the question of how the mathematical model and numerical algorithm used in the far field, including the boundary conditions at the external “artificial” boundary, affect the computed solution in the near field. A related question, which is perhaps as important, is how the numerical treatment of the far field affects the overall computational efficacy of the algorithm.

As shown by different authors (including the authors of the current manuscript) both theoretically and computationally, the proper treatment of artificial boundaries may have a profound impact on the overall accuracy and performance of numerical algorithms, as well as interpretation of the results, in many areas of scientific computing. It will therefore be natural to expect that molecular dynamics simulations as they pertain to studying the nano-structures are not exceptional. Consequently, the role and influence of the far-field boundary conditions have to be thoroughly investigated in this framework.

As a first stage of this investigation, we propose to carefully study the performance of the periodic boundary conditions, i.e., the current methodology of choice. For a given physical setup, numerical simulation with periodic boundary conditions will have to be run repeatedly for different locations of the far-field artificial boundary (from more remote to closer to the core of the computational domain). Subsequently, the computed solutions in the near field will have to be compared and their accuracy assessed for their dependency on the proximity of the outer boundary. Finally, one will have to decide how far away in the far field the artificial boundary needs to be placed so that the computed solution in the near field is essentially independent of the far field boundary. This will determine how expensive the numerical simulation will be. We expect that the proposed series of computational experiments will be rather demanding in computer time and memory. It will be crucial to combine the study of the boundary conditions with at least some of the possible techniques aimed at improving the overall numerical efficiency, see Sections 2, 4, and 7 of the current manuscript.

The next stage will be the mathematical and experimental testing of various types of far-field artificial boundary conditions. Again, a universal conclusion reached in the literature is that the accuracy and overall performance of such numerical techniques are determined primarily by whether or not the boundary conditions are capturing well the effects and solution properties in the truncated part of the original domain. Our experience in building and implementing special classes of highly-accurate local and nonlocal artificial boundary conditions for fluid flows and wave propagation (acoustics and electromagnetics) firmly corroborates this conclusion. In the framework of molecular dynamics as it applies to the simulation of nano-structures, we propose to look into the possibilities of constructing the artificial boundary conditions based on both kinetic (i.e., particle) and macroscopic (i.e., continuous medium) models employed in the far field. This is an extensive and far-reaching goal. In addition, if the approach to evaluating the forces based on first solving a differential equation for the potential of the force field (see Section 4), is adopted then the differential equation (e.g., the Poisson equation) will also need to be supplemented by artificial boundary conditions. Constructing highly-accurate artificial boundary conditions for the Poisson equation is a fairly well understood issue. These boundary conditions are known to perform particularly well when combined with a multigrid-based solver in the interior.

¹Consider, e.g., the surface of a cylinder — periodic boundary conditions obviously cannot be applied in the radial direction.

The methods for constructing the artificial boundary conditions vary from the more traditional and simple ones, such as those based on asymptotic approximations and reduced-dimension models in the vicinity of the artificial boundary, to more advanced techniques that require larger effort to implement but promise larger payoffs as well, such as the method of difference potentials. Other techniques include boundary integral equations coupled with the application of the fast summation rules in the spirit of those mentioned in Section 4. In the case of a macroscopic far-field model, the issues related to the transition from the kinetic description to the continuous medium (the so-called meso-scales) will also have to be identified and addressed to the appropriate level of detail (see the concluding paragraph of Section 2). This, of course, will be a separate task. Altogether, based on our previous experience, we expect that the benefits from the advanced treatment of outer boundaries will far outweigh the additional effort required for the development and testing of the corresponding algorithms.

6. Modeling. Usually, by modeling one means the simplified mathematical description of a rather complex physical or engineering system. Prime examples are, e.g.,

- The analysis of flutter of a wing with hanging engines is replaced by that of a system of lumped masses and a simple beam.
- The reduction of the full Navier-Stokes equations to the boundary-layer equations by means of neglecting the streamwise viscous terms, with the far field described by the Euler equations.
- The reduction of the description of an intense explosion to a set of self-similar solutions, achieved by simplifying the shock jump conditions to their asymptotic values.

In the first example, the modeling process is basically physical, i.e., physical, or rather engineering, considerations change the description of a complex structure to a simple one, albeit one that retains important characteristics of the original problem. In the other two instances (boundary layers and strong spherical shock waves) the mathematical description of the original system is known (i.e., via a set of non-linear partial differential equations and the accompanying initial and/or boundary conditions). The modeling in this case is done by simplifying the mathematics, usually through educated guesses or assumptions concerning the relative order of magnitudes of different terms in the equations, or in the boundary conditions. Thus the assumption that the second derivative of the velocity in the direction normal to the solid surface is much larger than that in the direction tangent to the surface leads one from the Navier-Stokes equation to the boundary-layer equations. These, in turn, in the steady-state admit self-similar solutions for certain geometries — something not possible with the original Navier-Stokes equations.

In the case of un-modeled problems, i.e., mathematical descriptions which were rigorously derived from the appropriate “laws of nature” (e.g., Newton’s laws when the physical scales are not too small and the speeds not too high), one usually does not have to worry about the mathematical validity of the system — it will be well-posed, and the solutions unique.² On the other hand, for the aforementioned two types of modeled systems the mathematical validity cannot be taken for granted. It is the task of the applied mathematician to ensure that the model retains the essential mathematical, as well as physical, features of the original system.

Since most problems in material sciences are quite complex, resorting to modeling is natural and appropriate. The initial stage of the modeling is probably best done by the material scientist who understands the physics and chemistry involved. (Although even at that stage it might be efficacious to involve an applied mathematician with a proper background. He may help resolve such issues as, for example, what is the

²In most cases, apparent non-uniqueness of such systems is due either to neglecting the second law of thermodynamics, or to the multiplicity of “steady states,” which is removed when the initial conditions are taken into account.

best way to simplify the description of a nano-tube embedded in a surrounding matrix, or, e.g., can and should one introduce “matching layers” to mediate between two force fields?) The contribution of the applied mathematician will probably be more significant in the second stage — that of trying to solve (usually by approximate/numerical methods) the mathematical problem as presented by the model. The experience gained from fluid mechanics, theory of elasticity, and electro-magnetics may be very useful in dealing with the new models presented by material science.

7. Sensitivity of MD results to truncation criteria for short-range forces. As has already been mentioned, MD simulations are computationally expensive. Part of the reason is the stiffness of the corresponding ODEs (see Section 2), the other one is related to the calculation of forces acting on all particles on every time step of the integration algorithm (see Section 4). Those forces are often subdivided into the long-range and short-range ones. In a system of N mutually interacting particles, a straightforward calculation of forces obviously results in an algorithm of complexity $\mathcal{O}(N^2)$. It is convenient to think of this algorithm as multiplication of the coefficient matrix of dimension $N \times N$ by the N -dimensional position vector for particles. The aforementioned matrix, which changes dynamically, is full if each particle interacts with all others. The corresponding $\mathcal{O}(N^2)$ cost becomes prohibitively expensive for large N 's. A possible remedy for long-range Coulomb forces is fast multipole summation schemes and/or multigrid type techniques (Section 4). The approach for short-range forces, e.g., those of Lennard-Jones type, is different. Namely, only the interactions between a given particle and its immediate neighbors are taken into account, whereas the forces from the particles that are further away are considered negligible and therefore disregarded.

To implement this idea efficiently, one needs to be able to determine at every time step, according to a pre-selected criterion, which particles are “close” to a given one, and which are “far away.” To minimize explicit calculation of distances between all particles, which would result in an $\mathcal{O}(N^2)$ effort anyway, one may build tracking algorithms based on creating and maintaining lists of close neighbors of all the particles involved. In this connection we mention that if, instead of considering all interactions, we start considering only those between the close neighbors, then the aforementioned coefficient matrix (on the right-hand side of the governing system of ODEs) instead of being full becomes sparse. The issues related to sparse matrices still constitute an active research area in numerical linear algebra. Therefore, we expect that efficiency of the MD simulations may benefit considerably from the implementation of the latest and future numerical algorithms for sparse matrices.

Another very important issue, which impacts on the previous paragraph, is choosing the criterion itself which defines the close neighbors of a given particle. Basically, this is the question of specifying a cut-off threshold for the distance, so that interactions with all particles beyond this range can be neglected (because the forces are short-range). To the best of our knowledge, current practices in MD simulations choose criteria for cut-off ranges mostly on ad hoc basis (e.g., the first neglected particle exerts only a given fraction of the force due to the nearest neighbor). This means that in principle we do not know the error due to disregarding the effect of all “far” particles. In our view it will be useful to search for cut-off criteria (most likely, of probabilistic nature) that keep the overall error due to neglecting “distant” particles below a prescribed level. Clearly, the analyses for constrained and unconstrained systems will differ.