THE EUROPEAN
PHYSICAL JOURNAL
SPECIAL TOPICS

Editorial

Editorial – Constraints: From physical principles to molecular simulations and beyond

G. Ciccotti[1,2,a] and P. Echenique[3,4,5,6,b]

[1] Dipartimento di Fisica, Università di Roma "La Sapienza", and CNISM, Piazzale A. Moro, 5, 00185 Rome, Italy
[2] School of Physics, University College Dublin, Room 302B, EMSC, Belfield, Dublin, Ireland
[3] Instituto de Química Física Rocasolano, CSIC, Serrano 119, 28006 Madrid, Spain
[4] Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Universidad de Zaragoza, Mariano Esquillor s/n, Edificio I+D, 50018 Zaragoza, Spain
[5] Departamento de Física Teórica, Universidad de Zaragoza, Pedro Cerbuna 12, 50009 Zaragoza, Spain
[6] Unidad Asociada IQFR-BIFI

The topic of constraints in molecular simulation is certainly not a new one: its origin dates back approximately to the 1970s and 1980s, not very long after systems of many bonded degrees of freedom were simulated for the first time using computers. This early introduction into the toolbox of computational physicists and chemists correlates well with their importance: it is not an exaggeration to say, for example, that a large fraction of molecular dynamics simulations performed nowadays includes them in one way or another. Through the years, this importance has motivated a large number of theoretical and numerical developments, producing a rich literature on the topic and solving many of the original caveats associated with the technique.

Some relevant modeling and technical problems however remain, whose resolution can potentially increase both the efficiency and accuracy of the calculations. Hence, research on constraints is still very active. A clear sign of this vitality is the fact that a number of groups all over the world are presently analyzing and revisiting the field, producing relevant innovations and insights. Many of these researchers met in September 2010 at the Zaragoza Center for Advanced Modeling (ZCAM), the Spanish node of the Centre Européen de Calcul Atomique et Moléculaire (CECAM), to participate in a vibrant workshop about the topic, and many of them are also authors in this special issue of The European Physical Journal.

The basic idea of imposing a constraint on an otherwise time-dependent function of the coordinates (and sometimes velocities) of the system is mostly applied with two clearly distinguished goals. First, constraints can be used to fix the fastest degrees of freedom of molecular systems (typically bond-length and bond-angle vibrations), with the objective of effectively reducing the dimensionality of the configuration space and, more importantly, of eliminating the highest frequencies from the

[a] e-mail: `giovanni.ciccotti@roma1.infn.it`
[b] e-mail: `echenique.p@gmail.com`

dynamics so that a larger time step can be used and the efficiency of the simulation increased. Second, constraints can be used to calculate free energy profiles along reaction coordinates. This use of constraints is conceptually very different but mathematically and technically very similar. While these applications do not exhaust the usefulness of constraints in modern simulations (think, for example, about their use in ab initio molecular dynamics), they are the focus of the chapters of this special issue. A number of important questions that are tackled (and sometimes partially answered) can be identified:

Theoretical, physical modeling questions:

- *Is there a sufficient justification for the introduction of (holonomic) constraints in molecular modeling? Computational interest aside, are classical stiff force fields more (or less) reliable than constraints? Would it be possible to give rigorous criteria to establish when constraints are justified and when not?*

  This is addressed from different angles in Chapter 1 where the canonical equilibrium of the holonomic and stiff model are compared and a unified formalism is provided to deal with them. In Chapter 5, a novel justification is provided for constrained models by coarse-graining the original unconstrained dynamics using projection techniques. Constrained models based on quantum mechanics are developed as a possible choice to assess the validity of classical approaches in Chapter 10, and the smoothing of the original differential equations is used to analyze and justify the imposition of constraints in Chapters 2 and 11.

- *Is the completely unconstrained situation the appropriate reference to assess the quality of constrained models (e.g., to answer the question above and the questions below) or should we use other alternatives like experimental data or quantum-based models?*

  This question is mostly left unanswered (in principle, there is no ontological reason pushing us to prefer springs to constraints in classical approaches), but it is discussed and different arguments are presented, especially in the chapters mentioned in the previous point.

- *Constraints can be defined by setting some internal coordinates such as bond lengths and bond angles to constant values (the most common approach), or they can be set to values that depend on the configuration of the molecule (this dependence can be extracted, for example, by a local minimization of the energy with respect to the hardest – to be constrained – coordinates at each configuration). The first option is termed "hard", and the second one "flexible". What is the more physically accurate way of introducing constraints: flexibly or non-flexibly? Does it make a big difference? Do flexible constraints allow one to constrain more degrees of freedom such as, say, bond angles, without altering too much the original, unconstrained physical model?*

  These questions are tackled in a number of chapters in this special issue. In Chapter 1, flexible constraints are shown to be more physically accurate at the level of equilibrium statistical mechanics, if the unconstrained case is used as a reference. In Chapter 2, an alternative to flexible constraints is presented which is based on correcting the potential energy function in such a way that hard constraints can be used, while still approximately recovering the properties of the flexible case.

- *Are constraints a sufficiently accurate tool for static properties? If so, under which conditions? Are they also good for dynamical properties? If so, under which conditions? If not, is there any alternative?*

The distinction between static and dynamic properties pervades the whole special issue, with some of the chapters dealing with only one aspect (equilibrium statistical mechanics), and others discussing both. Regarding the alternatives to constraints, they could come from two very interesting proposals in this special issue: corrected potentials (Chapter 2) and coarse-grained dynamics (Chapter 5).

- *Can constraints be understood as some sort of coarse graining of the original model? How is this idea related to the dynamical averaging of the fastest oscillations in the dynamics? Are coarse-graining approaches more reliable/efficient than (holonomic) constraints?*

  A very interesting first attempt to answer this question, as mentioned before, is presented in Chapter 5, where, instead of the constrained dynamics, the motion of the system is coarse grained through a projection technique that eliminates the fastest, hardest degrees of freedom in a rigorous way.

Computational questions:

- *How important are mass-metric tensor determinants and other differences between the holonomic and stiff models? How costly is it to compute them, in both the flexible and non-flexible cases?*

  These questions, first asked in the celebrated 1974 article by Fixman, are extensively reviewed and extended in Chapter 1, and also in many of the other chapters. Not surprisingly, the answer depends on the type of constraints imposed.

- *Standard Lagrange multiplier-based methods, which impose constraints on Lagrange equations of the first kind in numerical simulations, are the norm. What is the theoretical/computational value of alternative schemes, such as directly using generalized coordinates, or the impetus-striction formulation?*

  This question is explored in Chapter 7 of this special issue, where ambient space formulations of constraints are reviewed. In this type of formalism, the dynamics is not expressed using generalized coordinates, but instead in the whole space with Lagrange multipliers, and a certain freedom exists to choose the conjugate variable leading to the Hamiltonian form of the dynamics. In this chapter, ambient space formulations are also extended to the associated Langevin dynamics.

- *SHAKE and its developments are the standard numerical algorithms to impose constraints. Normally, they are not easy to parallelize. Is parallelization at all possible?*

  Chapter 9 discusses an interesting possible approach to parallelization.

- *Which are the more convenient ways of performing constrained molecular dynamics simulations in ensembles different from the microcanonical one?*

  This issue, important for practical applications of the computational modeling, is discussed in Chapters 6 and 7 (canonical ensemble) and in Chapter 8 (isothermal–isobaric ensemble).

Free energy calculations:

As for the second group of applications, many-atom systems display behaviors with typical times that are much larger than those attainable by brute force simulations (e.g., an average protein folds in seconds, while the accessible timescale is in the nanosecond range). The progress of these long-time behaviors can be parameterized

by collective functions of (many of) the degrees of freedom called *reaction coordinates*, and constraints can be used to fix these quantities at different points along the reaction. At each of these points, molecular dynamics simulations are performed, and the results are combined by a kind of thermodynamic integration to produce a free energy profile that would otherwise require an unfeasibly long time to compute. Various theoretical questions arise when using constraints in this context, and some of them are discussed in this issue. In Chapter 4, a thorough review of the techniques for computing free energy profiles based on constraints is provided, while in Chapter 3 two different definitions of the free energy associated with collective variables are discussed in a pedagogical style, clarifying some relevant issues.

Finally, let us acknowledge the collaboration of the many people who have made this special issue possible. First of all, the authors, whose rigor, good work and speed have been of course instrumental. Also, we are very grateful to the many anonymous referees, who, as is traditional in this "business", have performed the invaluable work of guaranteeing the quality and soundness of the articles (often leading to considerably improved final versions) without any form of personal benefit whatsoever, not even the tiny one of having their names mentioned. Thanks also to the staff of the Zaragoza Center for Advanced Modeling (ZCAM), especially to its Director, Michel Mareschal, and to the people of the Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), for their help in organizing the workshop where this special issue was born. Last but not least, we want to thank Christian Caron and Sabine Lehr at Springer for their invaluable kindness and professional skill.