Visual Analysis of Molecular Dynamics Data using Geometric and Topological Methods

Hans-Christian Hege

Zuse Institute Berlin (ZIB), Germany

(joint work with Norbert Lindow, Daniel Baum, Stefan Bojarovski, Vijay Natarajan)

Atoms and molecules are the constituents of matter, be it inorganic or organic, non-biological or biological. Thus all physical properties of matter can in principle be understood by analyzing these fundamental constituents and their mutual interactions.

However, these constituents are quantum mechanical objects, which leads to two problems: accurate numerical modeling requires a computational effort that often is infeasible and a comprehensible visual translation is often impossible, i.e. the objects remain unvisualized. From the mathematical point of view, the reduction of the quantum mechanical to a classical description entails non-trivial approximations. Nevertheless, decades of research have shown that (beyond situations where covalent bonds are formed or broken or the behavior depends sensitively on fine-tuned energy values) molecular phenomena can be well understood in the framework of the classical model. In this model the electrons follow adiabatically the classical nuclear motion and can be integrated out so that the nuclei evolve on a single Born-Oppenheimer potential energy surface (typically given by the electronic ground state), which in general is approximated in terms of few-body potentials. This is so-called classical molecular dynamics (MD), where the equations of motion of an N-atoms molecular system are solved in configurational space (3N Cartesian coordinates) and in momenta space (3N momentum variables), see e.g., [1]. Within well-known limits, MD has astonishingly good predictive power and allows to trace back the properties of a given system to a simple physical picture. It is the mental and mathematical model on which most of the molecular visualization is based on and which we also assume in the following.

In this framework the major factors determining molecular behavior are classical molecular forces and molecular shapes. They can be analyzed using MD simulations or experimental methods. The resulting data represent molecular configurations in spacetime that often are rather complex and difficult to comprehend. Therefore, molecular visualization has become an indispensable part of molecular analysis.

Various graphical representations of molecules have been invented and serve different needs in molecular sciences. However, aiming at visual analysis, i.e. at drawing physically correct conclusions from molecular depictions, are these representations suited? Important biophysical phenomena, like the transport of molecules through a biological membrane, often depend on fine details. It turns out that currently used types of molecular depictions sometimes do not allow us to draw correct conclusions and therefore more sophisticated representations are needed.

When trying to depict molecular details more accurately, a key question is: What is the real shape of a molecule? In contrast to macroscopic objects, the shape of a molecule cannot be determined with geometrical optics. An operational procedure is to determine the shape via repulsive forces that prevent interpenetration of bodies.
However, also in contrast to the macroscopic domain, no fine tip is available for ‘scanning’ the molecule. Instead, one can use only similarly sized objects, namely other atoms or molecules – which carry own force fields. The result therefore depends on the ‘probe’. Nevertheless, one gets exactly the crucial information, namely the mutual accessibility of molecules.

In the talk some recent developments will be reported that address two specific problems in molecular analysis: First, the gap between the fastest oscillations of covalent bonds of a molecule (some femtoseconds) and the time-scale of the interesting biological processes (some microseconds): this is bridged by computing the geometrically possible molecular paths in the presence of steric hindrances, which constrain the subset of dynamically possible, i.e. physical paths. Secondly, the uncovering and visualization of potential molecular pathways in complex, dynamic protein configurations. Specifically we will report on

- (a) a method for data-driven determination of effective atomic radii [5]
- (b) the computation of geometrically possible molecular paths within in cavities [4]
- (c) the interactive analysis of dynamical cavities and paths therein [2, 3]
- (d) the computation of realistic molecular surfaces that depict how close arbitrarily shaped molecules approach each other [6].

With these methods, for the first time, it is possible, e.g., to compute from MD trajectories the molecular paths within dynamic protein ensembles, leading to a better understanding, e.g. of molecular pumps in biomembranes.

REFERENCES