
Normal Mode Analysis and Propagators for Bio-Molecules.

Christopher R. Sweet
Computer Science and Engineering
University of Notre Dame.

Collaborators:

Jesús Izaguirre. Notre Dame.

Vijay Pande, Paula Petrone. Stanford University.



Normal Mode Analysis and Propagators

Normal Mode Analysis is based on a harmonic approximation of the dynamics of a system about an equilibrium position. Recent research suggests that conformational change can often be described by collections of the ‘low frequency’ modes. We consider the following:

- Normal Mode Analysis.
- Dynamics in the Normal Mode subspace.
- Time-coarsening in Normal Mode dynamics.
- Time-coarsening in the Normal Mode subspace using system forces.
- Conclusion.



Normal Mode Analysis

Molecular Dynamics can be described by Hamiltonian systems of the form $H = \frac{\mathbf{p}^T \mathbf{M}^{-1} \mathbf{p}}{2} + U(\mathbf{x})$, where the kinetic and potential terms are separable. For Harmonic systems which are purely harmonic we have the simplified potential term

$H = \frac{\mathbf{p}^T \mathbf{M}^{-1} \mathbf{p}}{2} + \frac{\mathbf{x}^T \Omega \mathbf{x}}{2}$, where Ω is the diagonal matrix of coefficients ω_i^2 . The Hessian of this potential energy is just Ω , since the bodies are not coupled. If we 'mass re-weight' this Hessian then the diagonal elements will be ω_i^2 / m_i , representing the frequencies of the oscillators.

Clearly the Hessian for the original system will not be diagonal due to the many body interactions.



Diagonalizing the Hessian

Given the Hessian A for the original system we can mass re-weight it by dividing each element by $\sqrt{m_i m_j}$. We can then diagonalize the matrix, with the Eigen Decomposition Theorem, such that $AQ = QD$, for matrix of eigenvalues D and matrix with eigenvector columns Q . If we chose the eigenvectors so that Q is orthonormal then we can map between the real and Normal Mode space easily.

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{M}^{-\frac{1}{2}} \mathbf{Q} \mathbf{c}, \quad \mathbf{c} = \mathbf{Q}^T \mathbf{M}^{\frac{1}{2}} (\mathbf{x} - \mathbf{x}_0)$$

where \mathbf{c} is the vector of mode positions, \mathbf{x}_0 the initial Conditions and \mathbf{M} the matrix of masses. We also have

$$\dot{\mathbf{x}} = \mathbf{M}^{-\frac{1}{2}} \mathbf{Q} \dot{\mathbf{c}}, \quad \dot{\mathbf{c}} = \mathbf{Q}^T \mathbf{M}^{\frac{1}{2}} \dot{\mathbf{x}}$$

The elements of D are the mode frequencies squared.

Normal Mode Equations of Motion

From either Hamiltonian or Newton we have equations of motion for our Normal Modes $\dot{c} = D\ddot{c}$, and from equipartition we expect that the average value of quadratic terms is $\frac{kT}{2}$ leading to the analytical solution at time τ

$$c_i = \sqrt{\frac{2kT}{d_i}} \sin(\sqrt{d_i}\tau)$$

Where d_i is the i^{th} diagonal of D . It has been observed that the frequency of the lower modes is inaccurate and equipartition overestimates the motion in these modes.

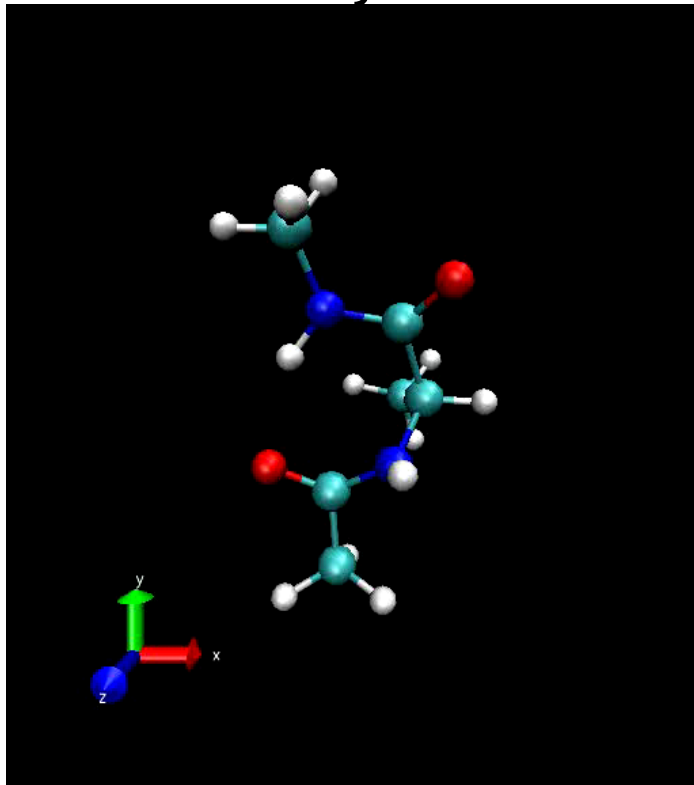
An alternative is to use the real system forces in the E.O.M

$$\ddot{c} = -Q^T M^{\frac{1}{2}} \nabla U \left(\mathbf{x}_0 + M^{-\frac{1}{2}} Q c \right).$$

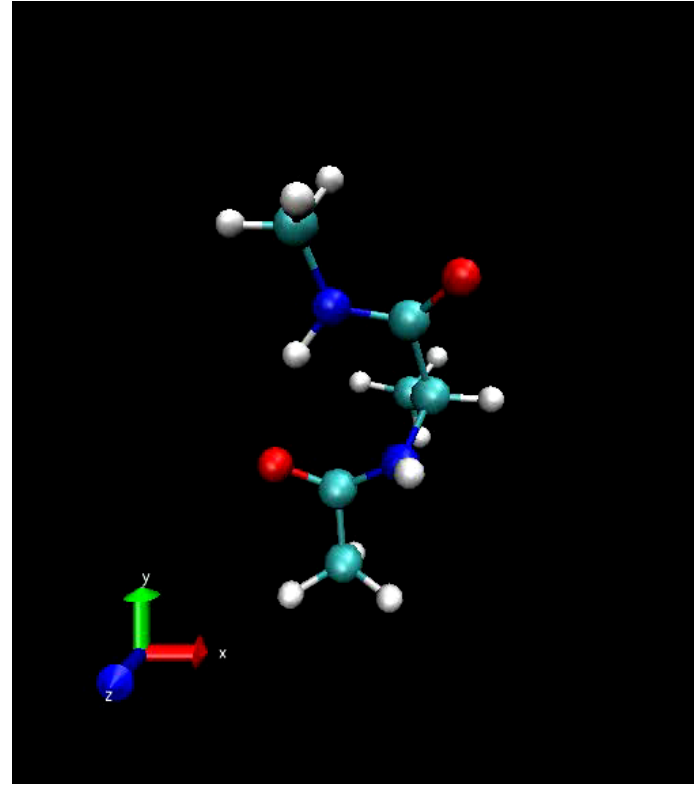
Dynamics in the Normal Mode subspaces

Using both the analytical and system force methods it is possible to propagate the system in the NM subspace.

Analytical



Real forces



Normal Mode Propagator artefacts

If we study the Normal Mode propagator with the analytical solution, we see some non-physical stretching of the bonds.

Clearly this solution does not respect forces and energies.

The real force propagator, by contrast, does respect these parameters. Even for the fast modes the real force propagator appears to give better results.

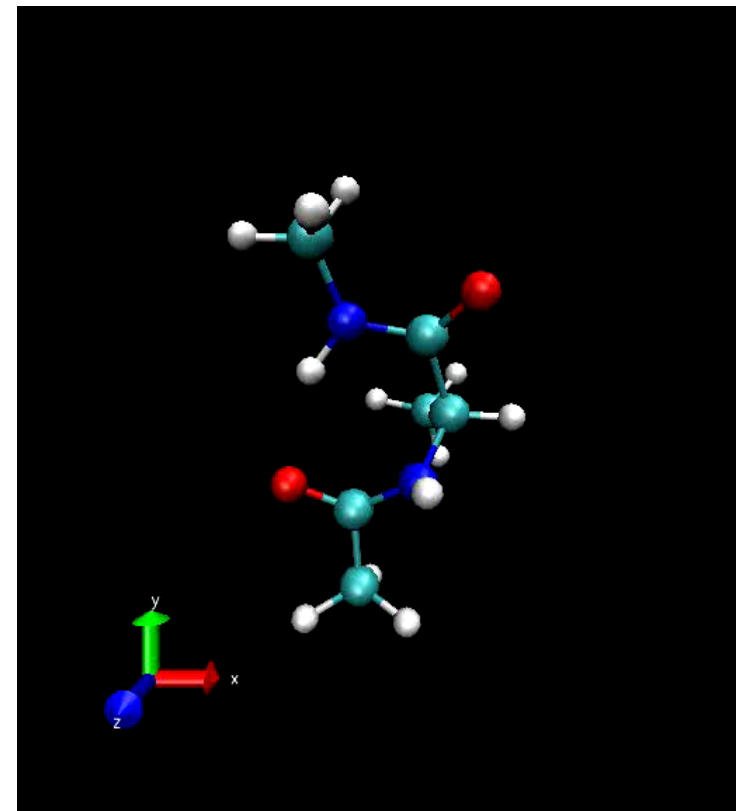
Time-coarsening in Normal Mode dynamics

Analytical Normal Mode propagators clearly offer huge computational savings as there are no forces to compute.

To use the real forces propagator has no computational advantage unless we can coarse grain in time by removing the fastest modes.

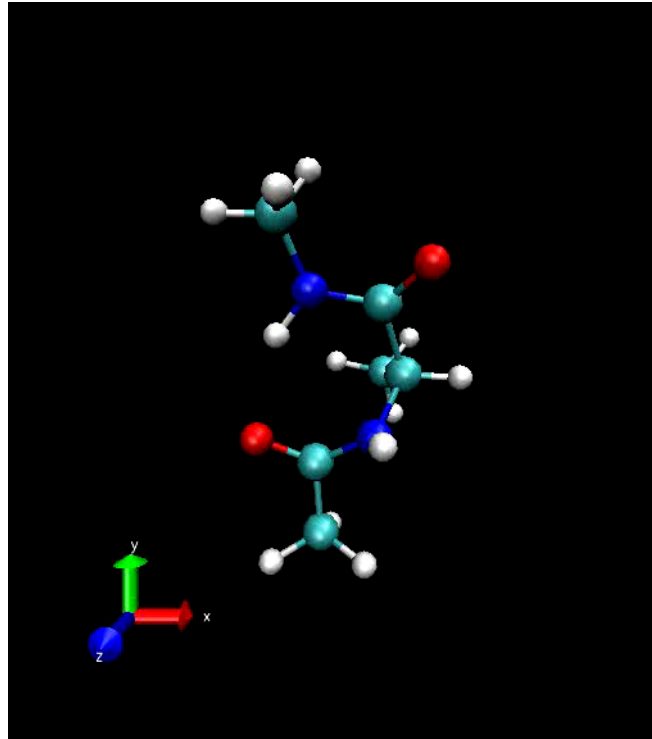
To remove modes for the analytical propagator merely requires that we fix their value at zero, since the modes are independent.

Analytical propagator
30 modes removed



Time-coarsening, real force propagator

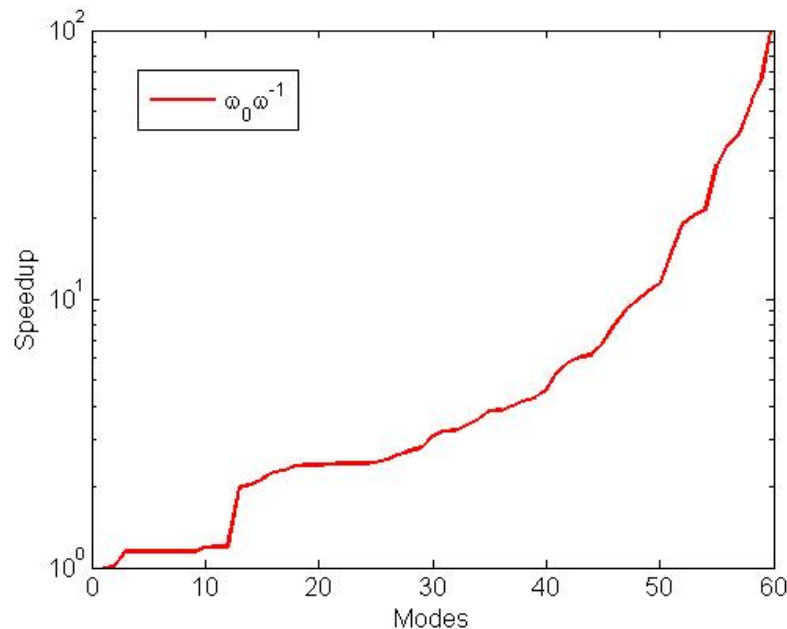
For the real force propagator, constraining modes is not as straightforward:



Even forcing a few fast modes to zero has catastrophic Results.

Time-coarsening with system forces

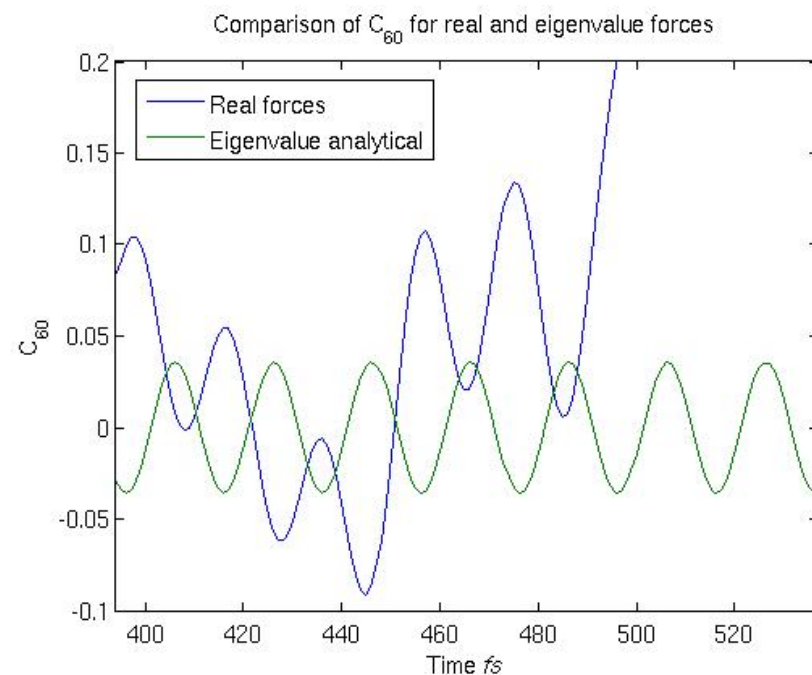
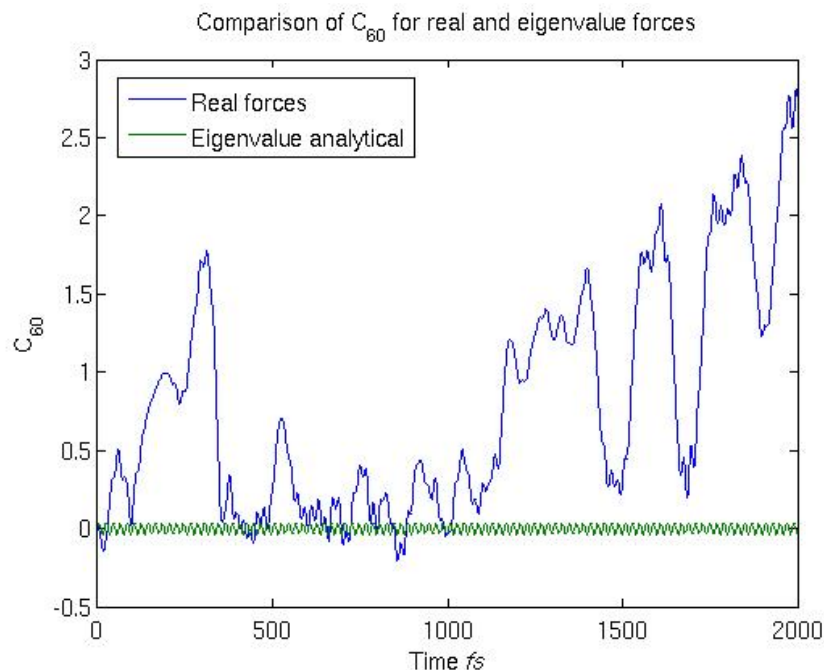
- For system force propagators there is no computational advantage over MD unless we can coarsen in time.
- Simple constraints on fast modes does not lead to stable propagation.
- Is coarse graining in time likely to yield computational savings?



Inverse frequency of modes when normalized to the fastest mode.

Normal Mode constraints

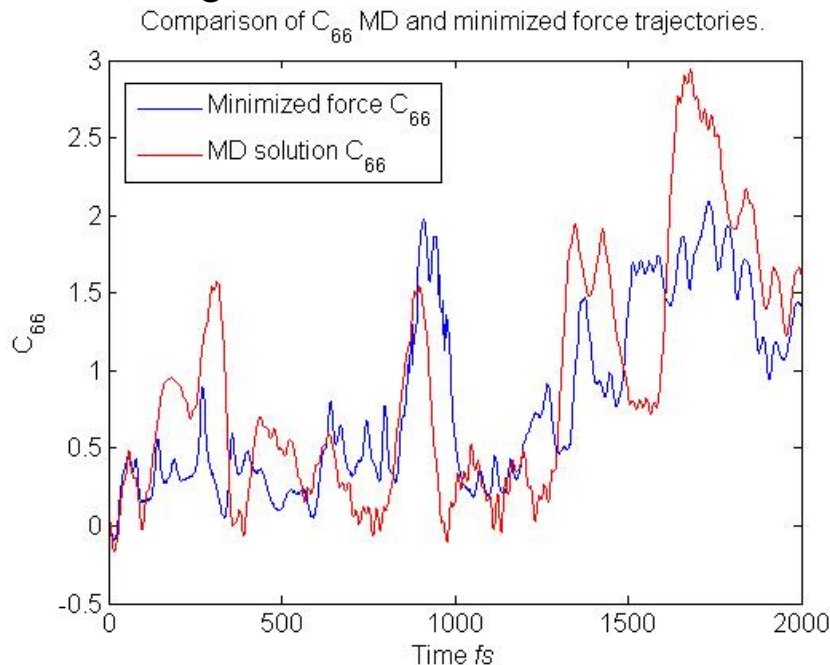
To constrain the fastest normal modes requires that we understand the instability introduced when modes are zeroed. The graphs represents the trajectory of the fastest mode, with no constraints.



Normal Mode constraints, zero force

For analytical propagators we have $\langle c_i \rangle = 0, \forall i$.

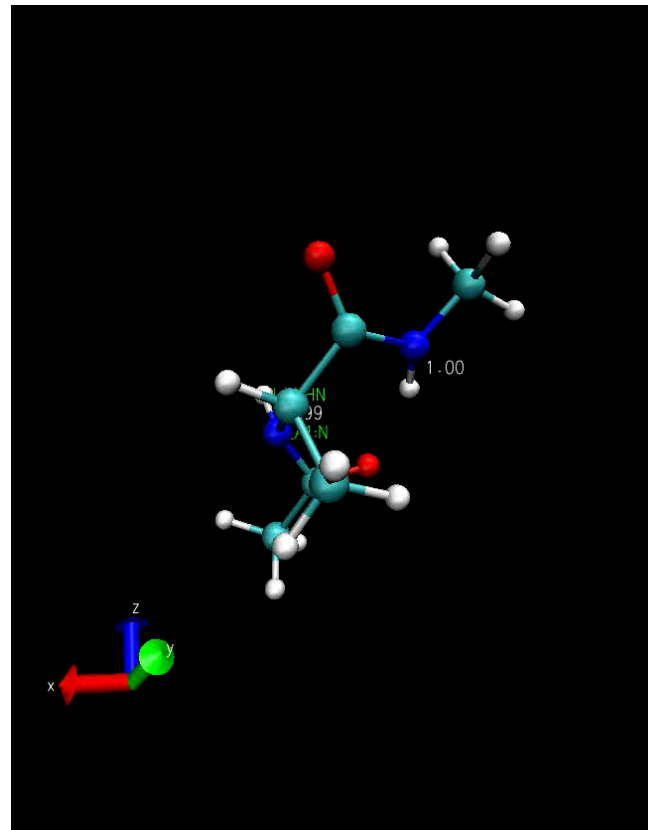
For the real force propagator we see that this is not the case, the c_i have a slowly varying component that modifies the initial \mathbf{x}_0 structure.



By minimizing the forces along the modes we can follow the trajectory without resolving the high frequency of the mode itself.

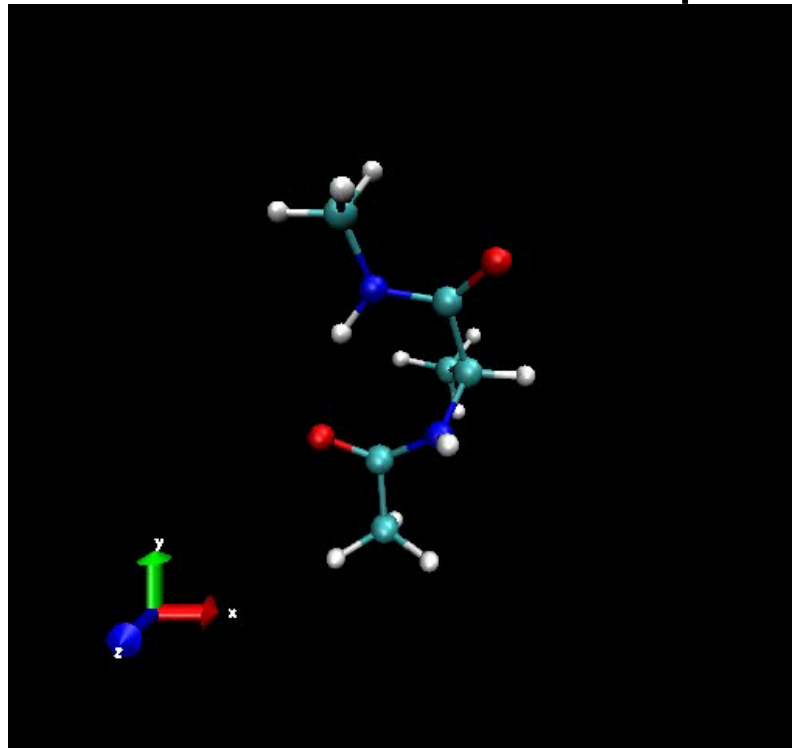
Normal Mode constraints, example

The following example constrains the fastest mode, N-H Bond for blocked Alanine Dipeptide, by minimizing the force along the mode.



Normal Mode, actual coarse graining in time

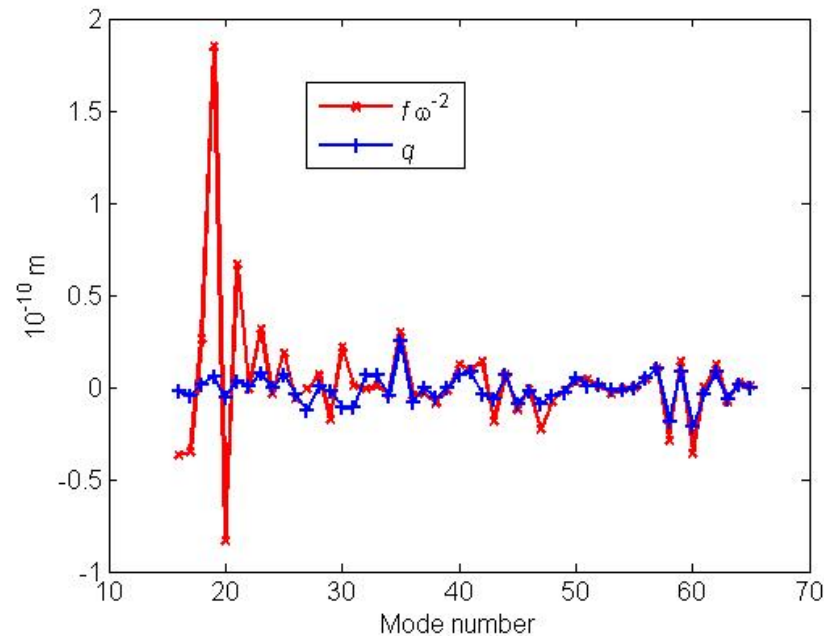
The following example constrains 13 modes on the Alanine Dipeptide. From the previous graph this should allow a step-size of 6fs which was achieved in practice.



Normal Mode, force minimizers

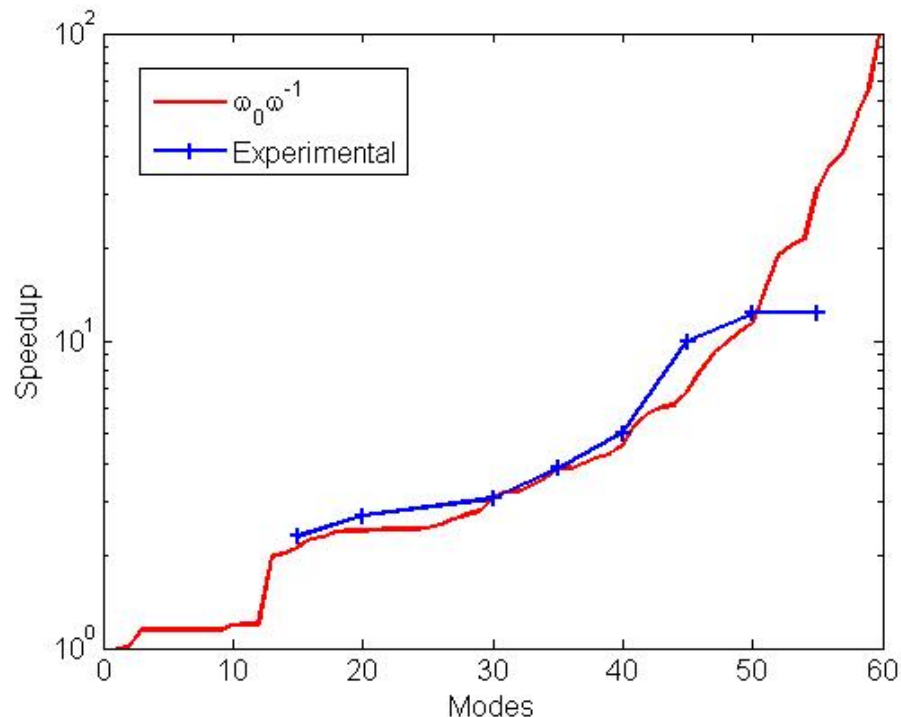
For the previous examples it was possible to use our knowledge of the eigenvalues as frequencies to provide an efficient force minimizer since force $f_i = d_i x_i$ (linear). As we consider more modes this model fails and different minimizers are required.

This leads us to general energy minimizers, such as conjugate gradient etc.



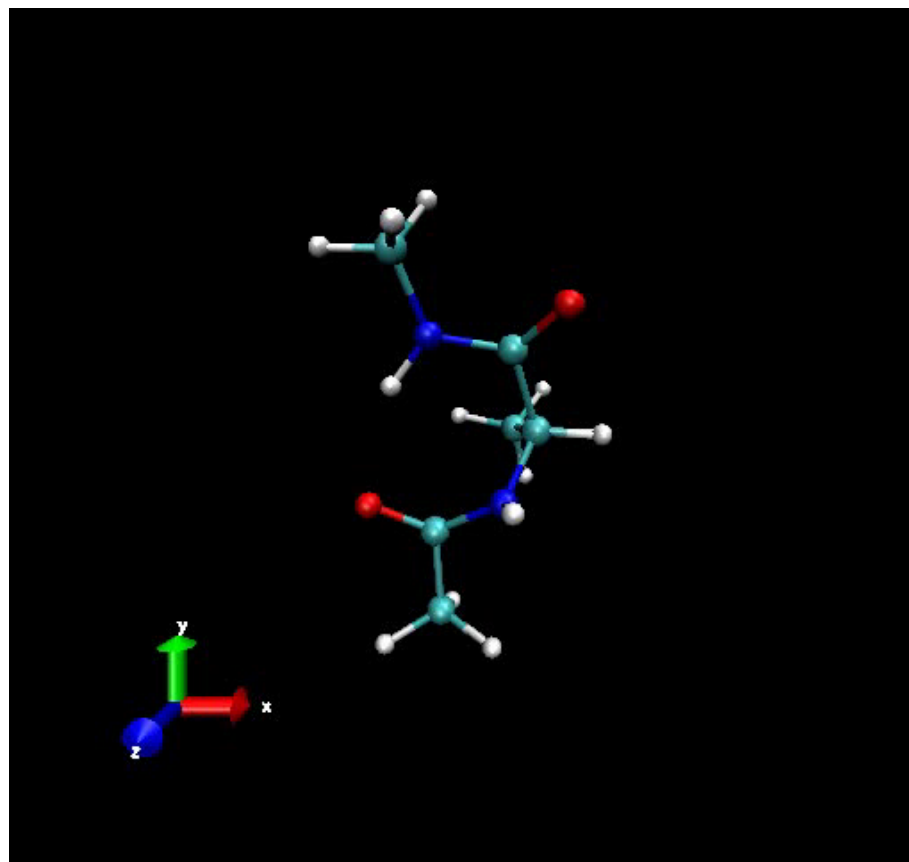
Normal Mode, estimate of time coarsening

The original estimate of potential time coarsening was based on the inverse frequency of the modes. The graph below compares this estimate to the results obtained.



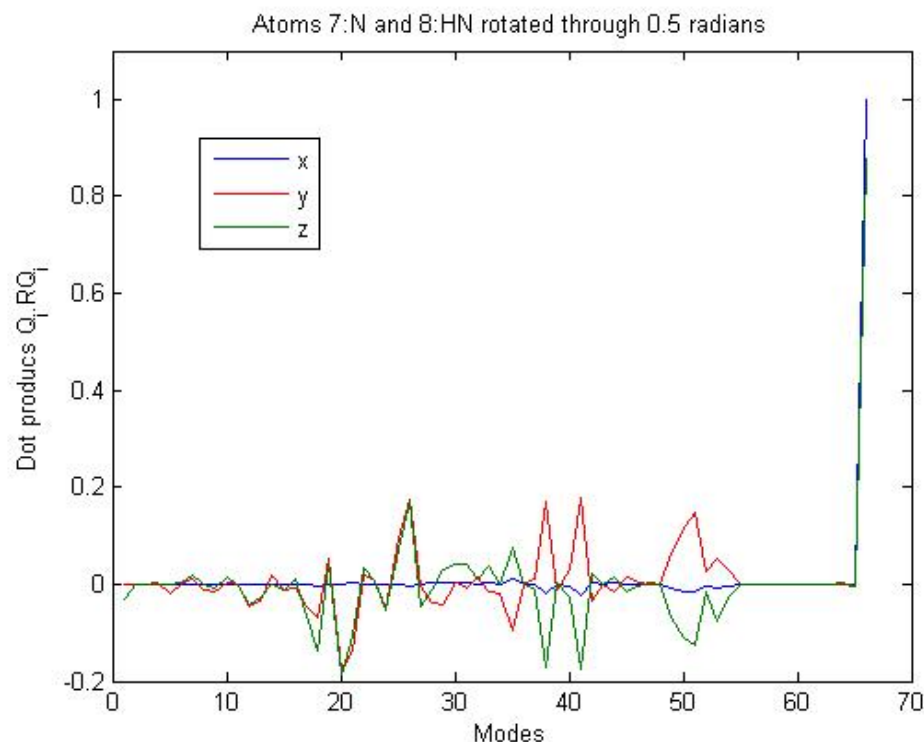
These results may be somewhat misleading. The propagator achieves something close to the theoretical maximum for the Verlet method, equivalent to H-bonds being resolved at a $2.5fs$ time-step, rather than $1fs$.

Normal Mode, 15fs time-step simulation



Does Minimization work?

The high frequency modes rotate through large angles, so is minimization accomplishing the constraints the we want?



The 'Inverse Problem'

If minimization is only minimizing in the complement subspace can we do it more effectively in MD space. We want to minimize the forces

$$\hat{f}_{\text{MD}} = f_{\text{MD}} - Qf_{\text{MS}} - \bar{Q}f_{\text{NS}} = (I - \hat{Q}\hat{Q}^{\text{T}})f_{\text{MD}}$$

since $(I - \hat{Q}\hat{Q}^{\text{T}})$ is idempotent we can use it in conjugate gradient schemes where $\Delta x = \hat{f}_{\text{MD}}^{n+1} + k\hat{f}_{\text{MD}}^n$. We can then update the positions with

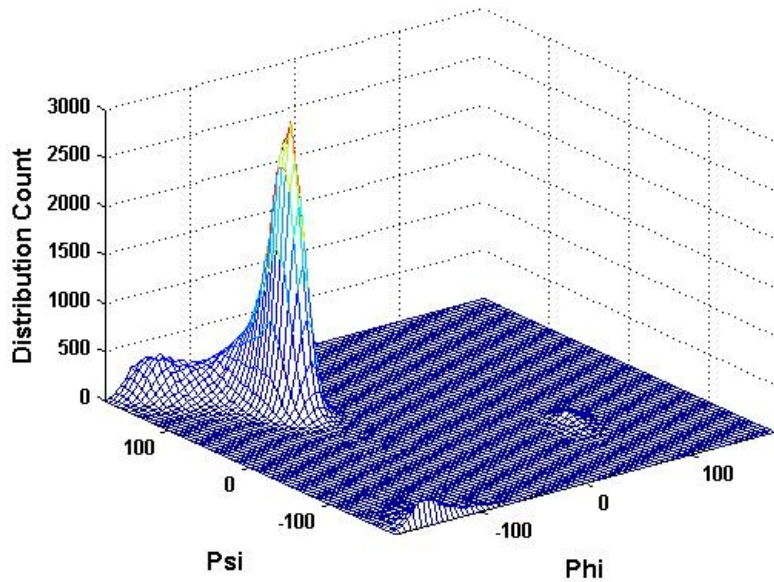
$$\hat{x} = (I - \hat{Q}\hat{Q}^{\text{T}})(x - x_0) + x_0 + M^{-\frac{1}{2}}Qc.$$

The Mode Space forces are

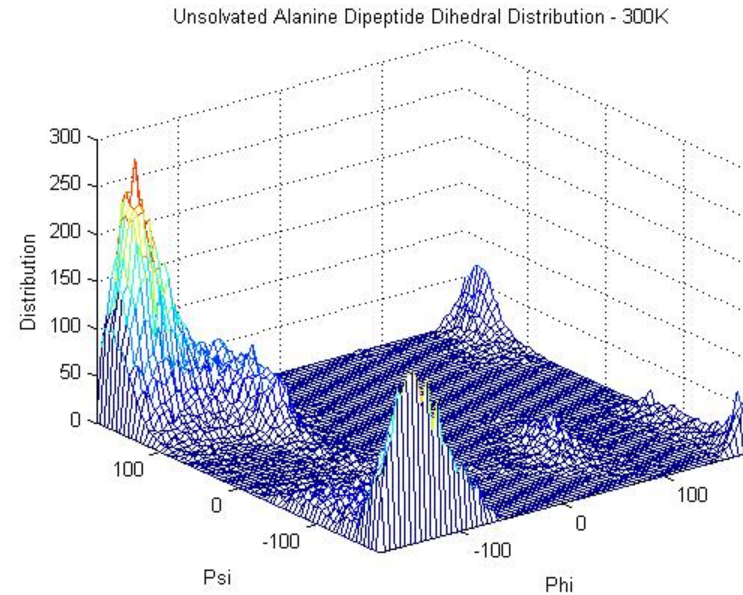
$$f_{\text{MS}} = Q^{\text{T}}f_{\text{MD}}.$$

Ramachandran plots for fixed modes.

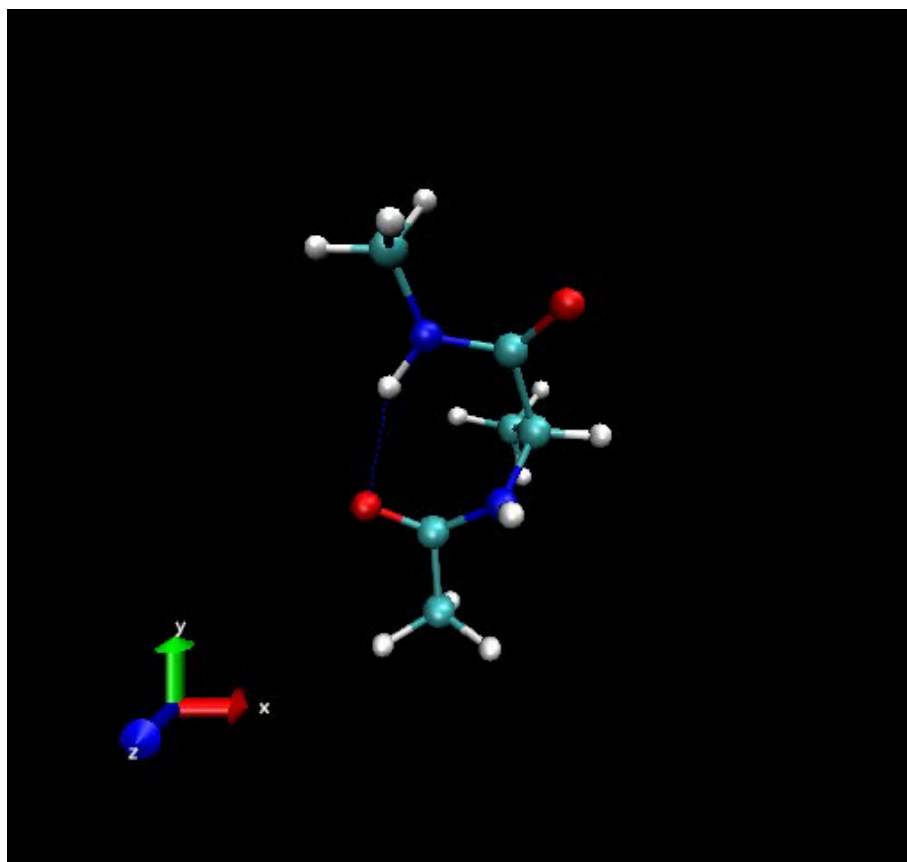
REM



45 fixed modes

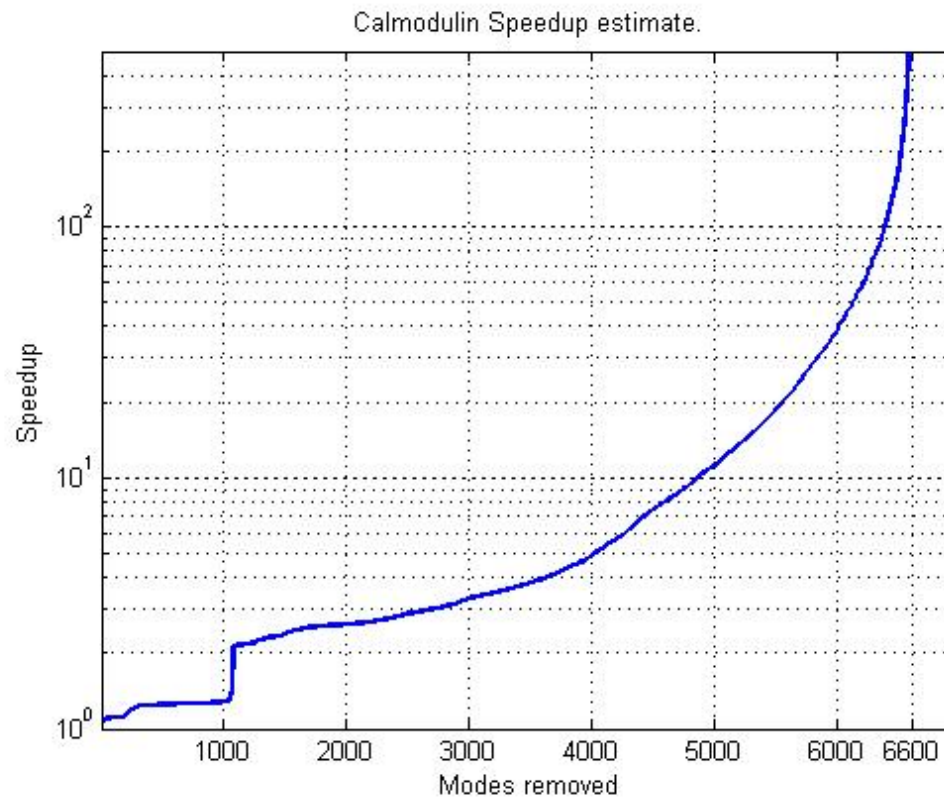


Sample Trajectory.

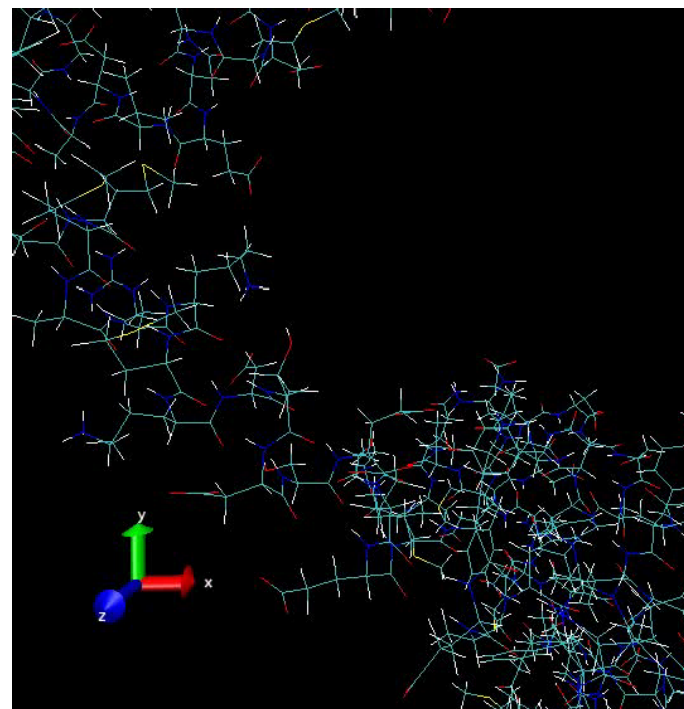
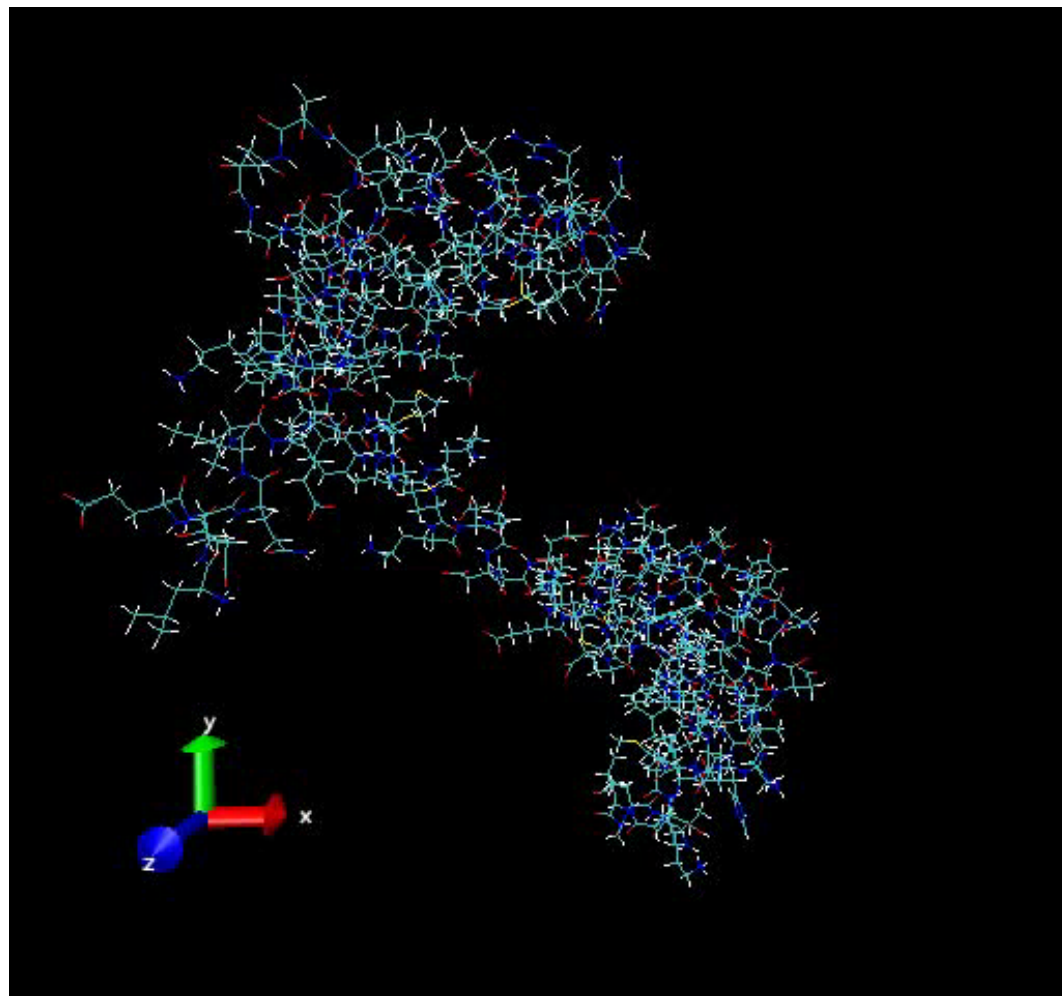


Real Biomolecules: Calmodulin, 2200 atom.

Potential speedup for Calmodulin:



Calmodulin simulation.



Zoomed showing
'constrained'
dynamics.

Map of Research.

- Dihedral/ Improper Hessian
- Langevin thermostating in Mode Space
- Scalability of minimizer
- Non-Hessian modes
- New Langevin method
- As part of REM
- Solvated alanine
- Calmodulin
 - Get initial configuration
- WW domain



Conclusions

- Even for small molecules the speedup is greater than 10 times.
- Real force Normal Mode propagators require minimizers to constrain modes.
- Langevin thermostating may be possible, not only for constant temperature and explicit solvent, but to replace the missing modes.
- Large bio-molecules should allow greater speedup due to the relatively higher number of low frequency modes.

