A Homotopy Method for Predicting the State of Minimal Energy for Chains of Charged Particles

Daniel M. Dunlavy
Applied Mathematics and Scientific Computation

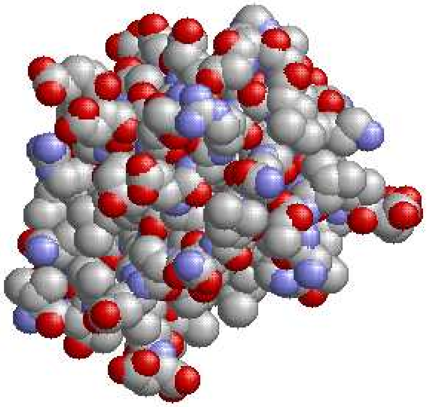
Acknowledgments

- Dianne O'Leary, *CS*
- Ron Unger, *UMIACS, Bar-Ilan University*

Introduction

- **Protein Folding**
  - Sequence of amino acids → three-dimensional structure
  - Minimum potential energy assumed for native structure

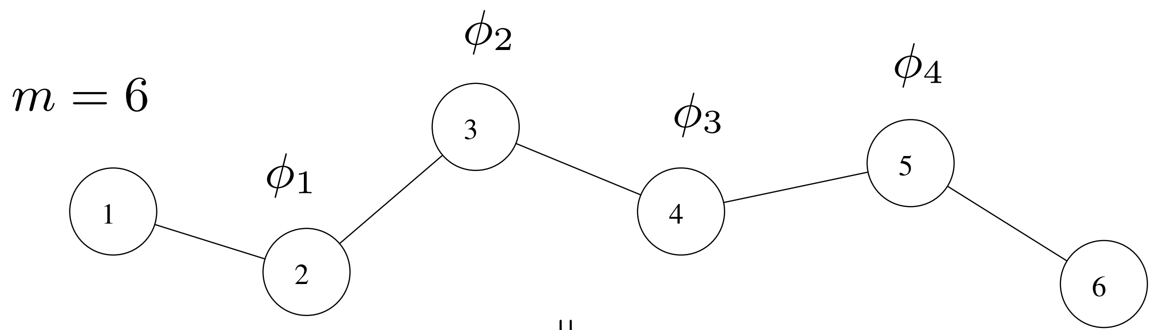
$\Longrightarrow$



- **Difficult Goal**
  - Find structure with minimum potential energy
  - Computationally intractable for large proteins

Formulation of the Problem

Chain of m charged particles with charges q_i (2D space)



$\Downarrow$

van der Waals Potential

$$E(\phi) = \sum_{i=1}^{m-2} \sum_{j=i+2}^{m} \left[\frac{q_i q_j}{R_{ij}(\phi)} + \varepsilon_{ij} \left(\left(\frac{\sigma_{ij}}{R_{ij}(\phi)} \right)^{12} - 2 \left(\frac{\sigma_{ij}}{R_{ij}(\phi)} \right)^{6} \right) \right]$$

$\Downarrow$

Optimization Problem

$$\min \quad E(\phi)$$

$$\text{s.t.} \quad 0 \leq \phi_i \leq 2\pi, \quad (i = 1, \ldots, m-2)$$

Solving the Optimization Problem

- **Difficulty**
  - Many local minima
  - Number of minima increases exponentially
- **Classic Approach**
  - Gradient methods (*e.g.*, steepest descent)
  - Good starting approximation needed
  - Converges to *local minimizer*
- **New Approach**
  - Homotopy method
  - Good starting approximation **not** needed
  - Improve likelihood of finding *global minimizer*

General Homotopy Method

Goal: Solve complicated nonlinear system which may have multiple solutions

$$f(x) = 0, \qquad (f : \mathbb{R}^n \rightarrow \mathbb{R}^n).$$

Steps to Solution:

Easy system: $e(x^0) = 0 \qquad (x^0 \;\; known)$

Homotopy: $h(x, \lambda) = \begin{cases} e(x) \,, & \lambda = 0 \\ f(x) \,, & \lambda = 1 \end{cases}$

$e.g.,\;\; h(x, \lambda) = \lambda \, f(x) + (1 - \lambda) \, e(x)$

Trace Path: Follow $h(x, \lambda) = 0$ from $\lambda = 0$ to $\lambda = 1$

Potential Energy Homotopy

Goal:

$$q^0 = [q_1^0, \ldots, q_m^0] \implies E^0(\phi)$$

$$q^* = [q_1^*, \ldots, q_m^*] \implies E^*(\phi)$$

$$\begin{array}{l} \min \ E^*(\phi) \\ \text{s.t.} \ \ 0 \le \phi \le 2\pi \end{array} \iff \left\{ \begin{array}{ll} \text{find} & \phi^* \\ \text{s.t.} & \nabla E^*(\phi^*) = 0 \\ & \nabla^2 E^*(\phi^*) > 0 \\ & 0 \le \phi^* \le 2\pi \end{array} \right.$$

Homotopy:

$$H(\phi, \lambda) = \nabla \left(\sum_{i=1}^{m-2} \sum_{j=i+2}^{m} \left[\frac{q_i(\lambda) q_j(\lambda)}{R_{ij}} + \varepsilon_{ij} \left(\left(\frac{\sigma_{ij}}{R_{ij}} \right)^{12} - 2 \left(\frac{\sigma_{ij}}{R_{ij}} \right)^{6} \right) \right] \right)$$

$$= \left\{ \begin{array}{ll} \nabla E^0(\phi) \, , & \lambda = 0 \\ \nabla E^*(\phi) \, , & \lambda = 1 \end{array} \right.$$

Tracing $H(\phi, \lambda) = 0$

$\phi^0 =$ global minimizer of $E^0(\phi)$

$\lambda_0 = 0$

$k = 0$

repeat until $\lambda_k = 1$

$$k = k + 1$$

$$\lambda_k = \lambda_{k-1} + (\Delta\lambda)_k$$

$$\phi^k \leftarrow \begin{cases} \text{using } \phi^{k-1} \text{ as initial guess} \\ \text{solve } H(\phi, \lambda_k) = 0 \end{cases}$$

end

$$\phi^* = \phi^k \quad \left[H(\phi^k, 1) = \nabla E^*(\phi^k) \approx 0\right]$$

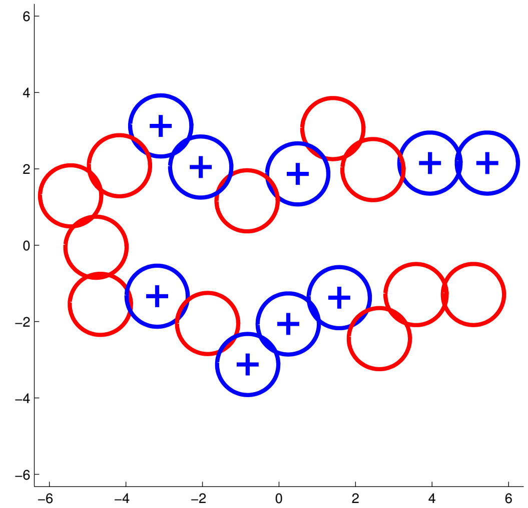
Example 1 – Negligible Difference

$m = 20$

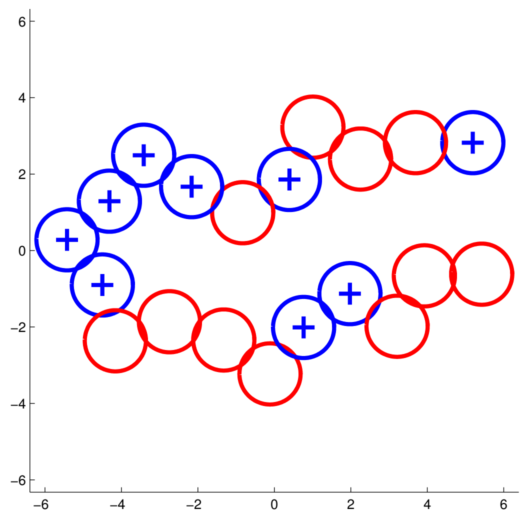
$q \in \{-1, +1\}$

$E^0(\phi) = -22.9708$

6 changes in q

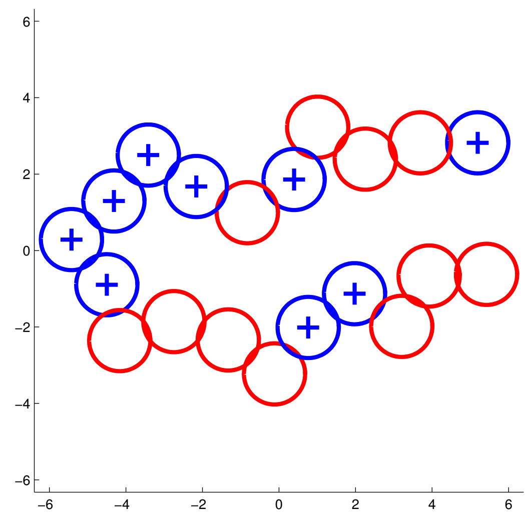


Gradient Method



$E^*(\phi) = -22.4510$

Homotopy Method



$E^*(\phi) = \boxed{-22.4511}$

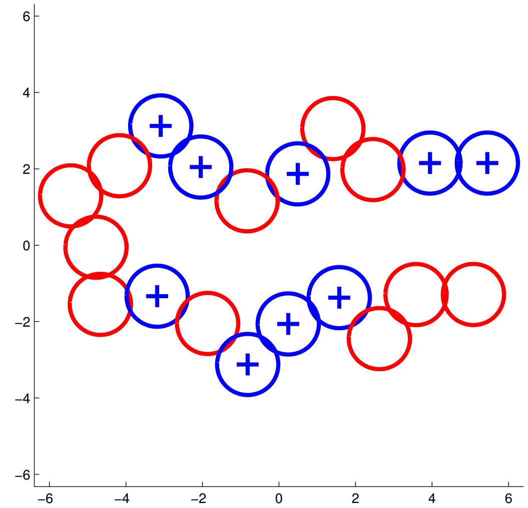
Example 2 – No Difference

$$m = 20$$

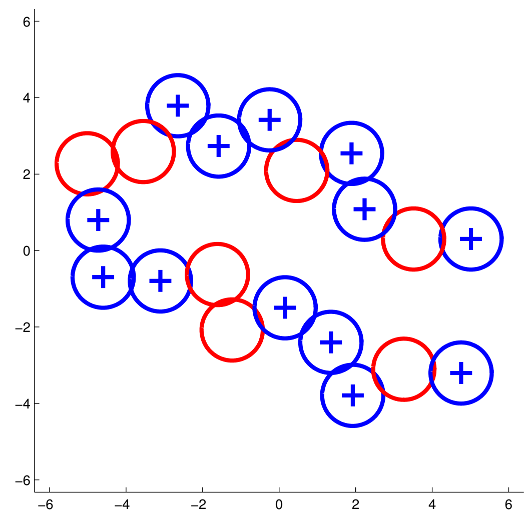
$$q \in \{-1, +1\}$$

$$E^0(\phi) = -22.9708$$

10 changes in q

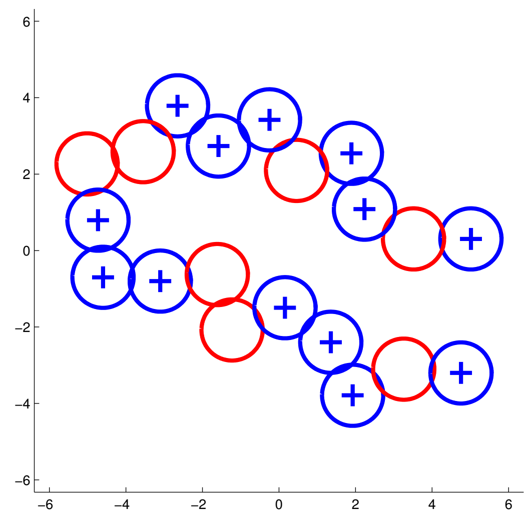


Gradient Method



$$E^*(\phi) = \boxed{-20.0044}$$

Homotopy Method



$$E^*(\phi) = \boxed{-20.0044}$$

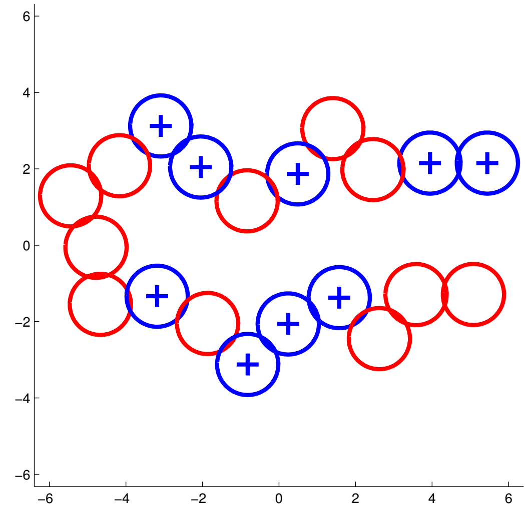
Example 3 - Qualitative Difference

$$m = 20$$

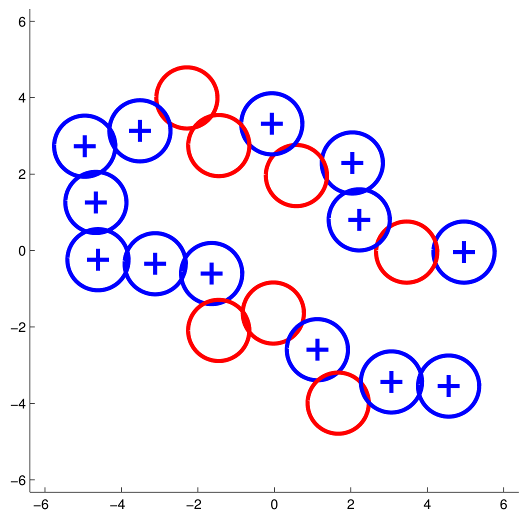
$$q \in \{-1, +1\}$$

$$E^0(\phi) = -22.9708$$

16 changes in q

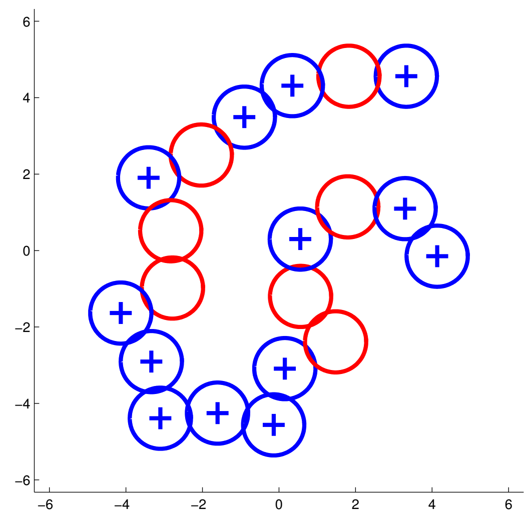


Gradient Method



$$E^*(\phi) = \boxed{-18.8808}$$

Homotopy Method



$$E^*(\phi) = \boxed{-19.4268}$$

Conclusions

- **Homotopy Method**
  - Rivals gradient methods (GM) in accuracy
  - Outperforms GM when many charges change
  - More function evaluations than GM

Future Work

- **Extend ideas to proteins**
- **Software interface for energy models**
  - Models: 2D/3D, Thirumalai/Klimov, AMBER
- **Software interface to Protein Data Bank (PDB)**
  - PDB $\rightarrow$ internal format $\rightarrow$ PDB
- **Homotopy tracing algorithm**
  - Predictor-corrector: ODE + optimization
  - ODE solvers: Euler, Runge-Kutta, etc.
- **Validation of the method**
  - Use existing PDB data