Efficient Geometry Optimization for Molecular Clusters

Suzanne M. Shontz
Department of Computer Science and Engineering
University of Minnesota
shontz@cs.umn.edu

Workshop on Advanced Electronic Materials
July 29, 2005

My Academic Background

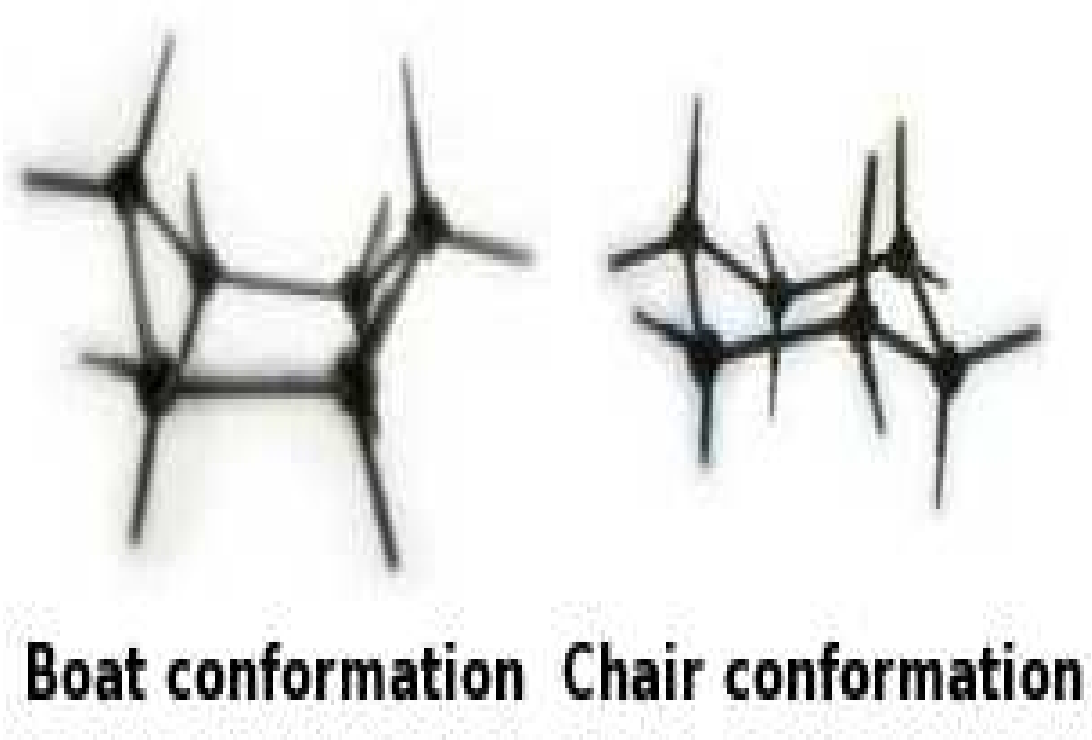
- B.A. in Mathematics, B.S. in Chemistry
  (University of Northern Iowa, 1999)
- M.S. in Computer Science, M.S. in Applied Mathematics
  (Cornell University, 2002)
- Ph.D. in Applied Mathematics (Cornell University, Jan. 2005)
- Postdoctoral Associate in Computer Science
  (University of Minnesota, 2004-on)

The PARSEC Package

- PARSEC = Pseudopotential Algorithms for Real Space Energy Calculations
- The goal of PARSEC is to solve the electronic structure of confined systems.
- Two-step approach:
  1. Electronic structure reconstruction: The potentials and charge densities must be self-consistent
  2. Geometry optimization/structural relaxation: Find the lowest energy conformation
- This is an iterative process.

A Motivational Example

- There are two main cyclohexane conformations.

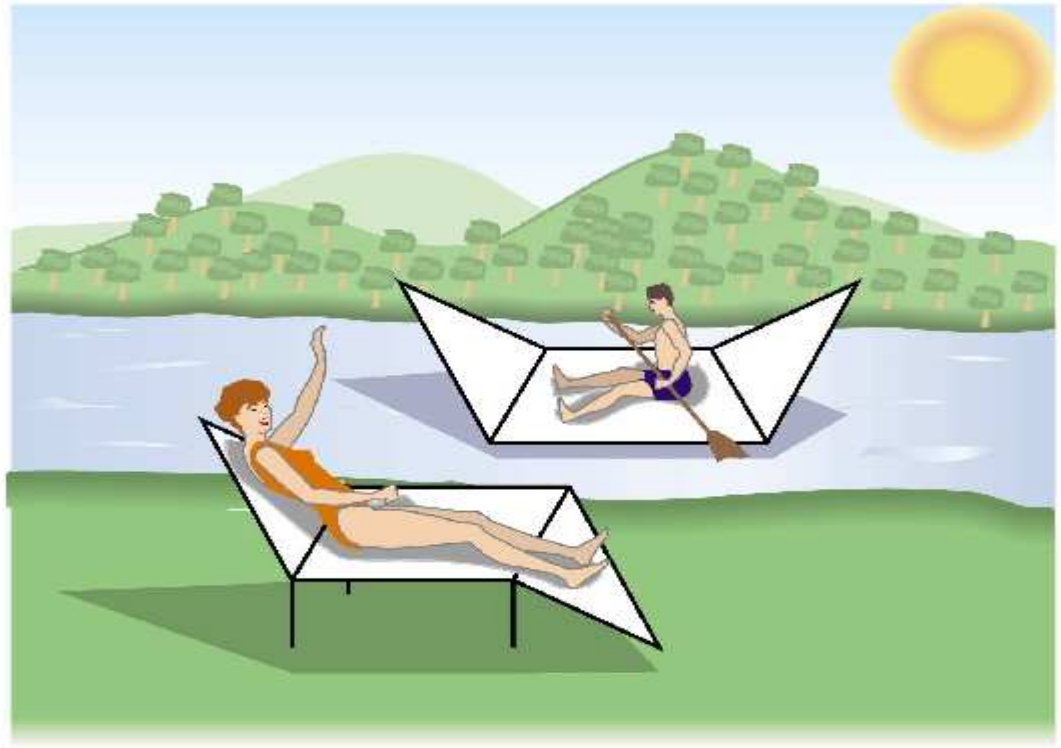


From http://en.wikibooks.org

- Which one occurs in nature? How do we figure this out?

A Simpler Problem!

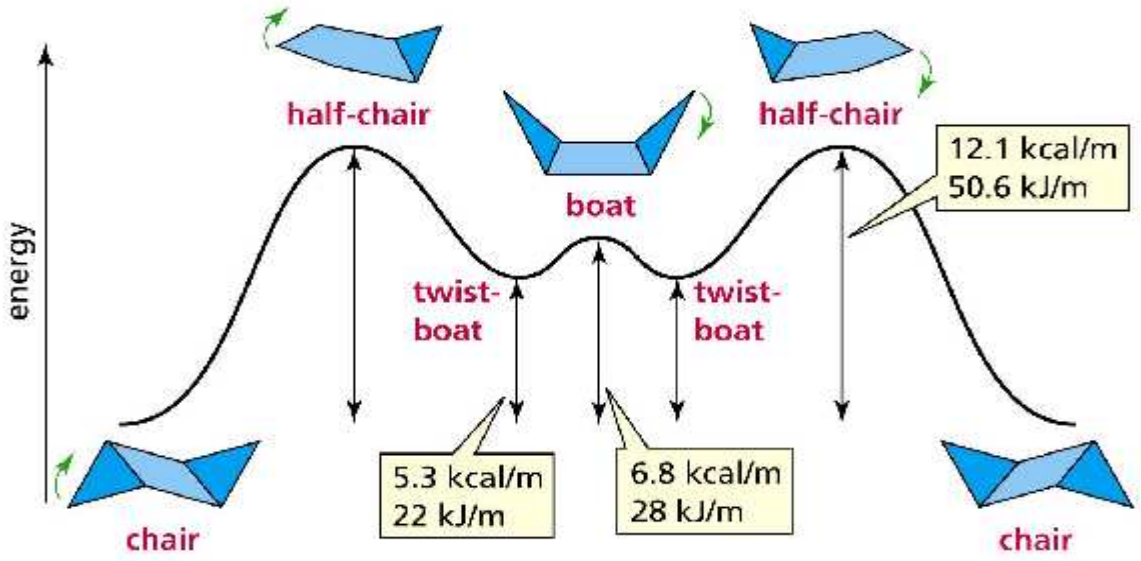
- Put another way - who is using less energy?!?



From Organic Chemistry by Paula Yurkanis Bruice at http://wps.prenhall.com.

Potential Energy Minimization

- We can determine which conformation occurs in nature by minimizing the potential energy.



From Organic Chemistry by Paula Yurkanis Bruice at http://wps.prenhall.com.

The Geometry Optimization Problem

- The goal is to minimize the potential energy of the molecular structure.
- We do this by repositioning the atoms based on a fixed charge density (at a given step).
- This is a nonlinear optimization problem.
- Here $f(x) =$ potential energy, $\nabla f(x) =$ forces.

Characteristics of the Optimization Problem

- The objective function is highly nonlinear.
- Expensive function, gradient evaluations: evaluating $f(x)$ means a self-consistent loop calculation
- The objective function and gradient are noisy.
- The Hessian, i.e., $\nabla^2 f(x)$ is not available.
- There may be many local minima.

Algorithms for Unconstained Minimization

Newton's Method: A First Approach

- Given $f : \mathbb{R}^n \to \mathbb{R}$, twice continuously differentiable, $x_0 \in \mathbb{R}^n$; at each iteration k,

$$\begin{aligned} \text{solve} \quad \nabla^2 f(x_k)\, s_k^N &= -\nabla f(x_k), \\ x_{k+1} &= x_k + s_k^N. \end{aligned}$$

- This attempts to find a zero of $\nabla f(x)$ by minimizing $m_k(x_k + s_k) = f(x_k) + \nabla f(x_k)^T s_k + \frac{1}{2} {s_k}^T \nabla^2 f(x_k) s_k$.

- One Issue: PARSEC doesn't provide analytic Hessians.
  - Impossible to get via automatic differentiation
  - Too expensive to evaluate via finite-differences

Quasi-Newton Methods: A Revised Approach

- Given $f : \mathbb{R}^n \to \mathbb{R}$, twice continuously differentiable, $x_0 \in \mathbb{R}^n$; at each iteration k,

$$\begin{aligned} \text{solve} \quad B_k \, s_k^N &= -\nabla f(x_k), \\ x_{k+1} &= x_k + s_k^N, \end{aligned}$$

  where $B_k \approx \nabla^2 f(x_k)$.

- Issue: How to approximate B_k?
  - Several types of approximations
  - Hessian updates are popular

Descent Directions

- Natural global strategy for minimization: make sure that each step decreases the value of f.
- Choose direction p from x_k such that $f(x_+) < f(x_k)$, where $x_+ = x_k + \alpha p$ for some $\alpha \in \mathbb{R}$. This is a descent direction.
- Mathematically, this happens when $\nabla f(x_k)^T p < 0$.
- If B_k is symmetric positive definite, the quasi-Newton step is a descent direction.

Secant Updates

- The secant equation in 1D is $B_+ = \frac{f'(x_+) - f'(x_k)}{x_+ - x_k}$.
- In general, B_+ should satisfy $B_+ \; s_k = y_k$, where $s_k = x_+ - x_k$ and $y_k = \nabla f(x_+) - \nabla f(x_k)$.
- The BFGS Update (Broyden, Fletcher, Goldfarb, Shanno):

$$B_+ = B_k + \frac{y_k {y_k}^T}{{y_k}^T s_k} - \frac{B_k s_k {s_k}^T B_k}{{s_k}^T B_k s_k}.$$

  - This is a symmetric, positive definite, rank-two secant update.

Secant Updates (continued)

- The Symmetric Rank-One (SR1) Update

$$B_+ = B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k}$$

  - This secant update is NOT positive definite.

- Combined Hessian Update (Farkas and Schlegel)

$$B_+ = B_k + \phi\ \Delta {B_k}^{SR1} + (1-\phi)\Delta {B_k}^{BFGS},$$

where $\phi = \sqrt{\phi^{Bofill}} = \sqrt{\frac{(y_k - B_k s_k)^T s_k (y_k - B_k s_k)^T s_k}{(y_k - B_k s_k)^T s_k {s_k}^T (y_k - B_k s_k) {s_k}^T s_k}}$.

  - This update combines the best properties of each.

Line-search Methods

- Idea of a line-search algorithm: Given a descent direction p_k, take a step in that direction that yields an "acceptable" x_{k+1}.
- At iteration k:
  1. Calculate a descent direction p_k,
  2. Set $x_{k+1} = x_k + \lambda_k p_k$ for some $\lambda_k > 0$ that makes x_{k+1} an acceptable next iterate.
- Step acceptance rules are necessary for obtaining:
  - decrease in $f(x)$
  - appropriate steplengths.
- Choose shorter steps in the quasi-Newton direction.

Line-searches and Descent Directions

This illustrates a descent direction within a line-search method.

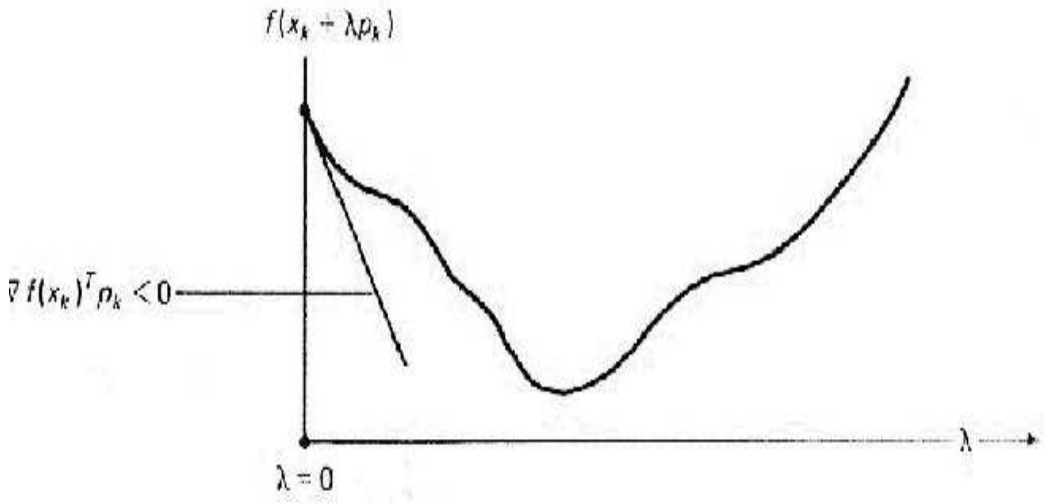


Figure 6.3.1 A cross section of $f(x)$: $\mathbb{R}^n \to \mathbb{R}$ from x_k in the direction p_k

From Numerical Methods for Unconstrained Optimization and Nonlinear Equations by Dennis and Schnabel

Note the importance of taking steps of appropriate length!

Backtracking Line-search Framework

Given $\alpha \in (0, \frac{1}{2}), \quad 0 < l < u < 1$

$\lambda_k = 1;$

while $f(x_k + \lambda_k d_k) > f(x_k) + \alpha \lambda_k \nabla f(x_k)^T d_k,$ do

$\lambda_k = \rho_k \lambda_k$ for some $\rho_k \in [l, u];$

$x_{k+1} = x_k + \lambda_k d_k;$

Model-Trust Region Methods

- Line-search = take a shorter step in the quasi-Newton direction.
- Trust region = shorten steplength and choose the direction.
- Need to have an idea of maximum allowable steplength, δ_k.
- Choose direction by solving:
  $\min\ m_k(x_k + s) = f(x_k) + \nabla f(x_k)^T s + \frac{1}{2} s^T B_k s,$
  subject to $\|s\|_2 \leq \delta_k$.
- δ_k provides a region in which we can trust m_k to adequately model f.

Trust-region Method: Pictorial Version

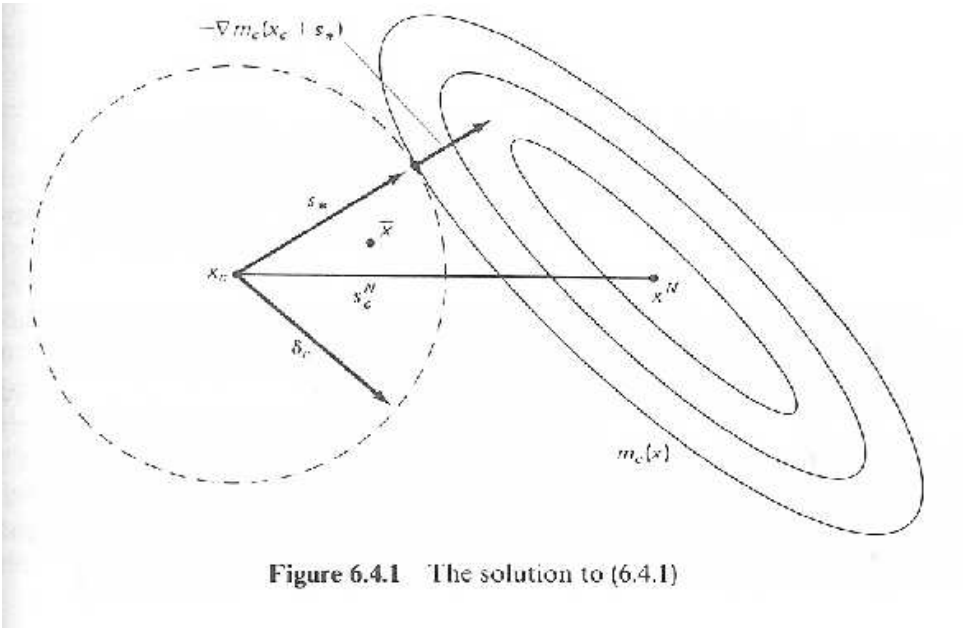


Figure 6.4.1 The solution to (6.4.1)

From Numerical Methods for Unconstrained Optimization and Nonlinear Equations by Dennis and Schnabel

The solution is either the Newton step (if it lies within the trust region) or a modified-Newton step (of length δ_k).

A Global Step of the Trust-Region Method

Given $f : \mathbb{R}^n \to \mathbb{R}$, $\delta_k \in \mathbb{R}^n$, $H_k \in \mathbb{R}^{n \times n}$ symmetric and positive definite:

repeat

(1) $s_k := \text{approximate solution to trust} - \text{region problem}$

$x_+ := x_k + s_k,$

(2) decide whether x_+ is acceptable, and calculate a new δ_k.

until x_+ is an acceptable next point;

$\delta_+ := \delta_k.$

Tensor Methods

- Can be thought of as an extension of Newton's Method
- Bases each iteration upon a fourth order model of the objective function:

$$m_T(x_+) = f(x_k) + \nabla f(x_k) \cdot d + \frac{1}{2}\nabla^2 f(x_k) \cdot d^2 + \frac{1}{6} T_k \cdot d^3 + \frac{1}{24} V_k \cdot d^4$$

- T_k and V_k are low-rank third and fourth order terms that cause the model to interpolate the already calculated function and gradient of the previous iterate.
- The main difficulty is in determining an accurate Hessian.

Conclusions

1. It is important to minimize the energy of the molecular system by optimizing the positions of the atoms.

2. This is a difficult nonlinear optimization problem.

3. Quasi-Newton and tensor methods form two categories of approaches to finding a local minimum.

4. Finding the global minimum is a more difficult problem. For now, we will be content with finding a local minimum.

Acknowledgements

People:

- Yousef Saad, Jim Chelikowsky
- Murilo Tiago - Parsec consultation

Funding:

- NSF/DMR-0325218