3805ICT - Assignment 2

Damon Murdoch (s2970548)

July 27, 2018

1 Introduction

The molecular model problem is based around the concept of 'N' atoms connected by rigid bonds of unit length constrained in two dimensional space. The potential energy of such a system, 'V' is given by the scaled, pairwise addition of Lennard-Jones potentials. The purpose of this investigation is to develop an algorithm for minimising the value of 'V' for a molecular model for any given 'N' number of atoms. The algorithm will be implemented using the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm to minimise the energy in the system iteratively, using the gradient of the change in energy for each atom at each step to minimise the energy cost of the system. While this problem operates on an extremely simple representation of an atom in two dimensional space, the investigation is useful for developing molecular structure optimisation algorithms as the minimum energy for the system can be determined for any dimension without undue complexity. The following report will provide a detailed description of the problem, the algorithm implemented to solve the problem and the resulting test data provided by the investigation.

2 Literature Review

A molecule is defined as "the simplest unit of a chemical substance, usually a group of two or more atoms "(Dictionary, 2018). an atom is defined as "the smallest unit of any chemical element, consisting of a positive nucleus surrounded by negative electrons"(Dictionary, 2018). Understanding the stable configurations of a molecule is important because it enables us to understand its properties and behavior with respect to its structure. When a molecule is constructed using computational chemistry software it may not be given a stable initial state. This means that if the same molecule were to be constructed in reality, it may be unstable and behave unexpectedly. As a result of this energy minimisation, also known as geometry optimisation is performed to find a stable solution state for the molecule.

Energy Minimisation is a numerical procedure which is used to find the minimum potential energy of a molecular system starting from a higher energy initial structure. During this process, the geometry of the molecule is adjusted iteratively such that the overall potential energy of the system is reduced. The search is terminated after a set number of iterations are performed or an accepted global minimum point is reached. CASTEP is a quantem mechanics based program designed for solid-state materials science, and can be utilised for geometry optimisation using BFGS and damped molecular dynamics. The BFGS minimiser is used primarily as it has the ability to perform cell optimisation, including optimisation at fixed external stress levels. This implementation involves a Hessian model in the mixed space of internal and cell degrees of freedom, in order to ensure both lattice parameters and atomic coordinates can be optimised. While left off by default, constraints can be applied to the model such as the fixing of atom positions, or fractional coordinates of the atoms.

The BFGS algorithm for numerical optimisation is used to minimise a function f(x), where 'x' is a vector in R^n, and 'f' is a differentiable scalar function. There are no constraints enforced on the value of 'x' by the algorithm. It operates by taking an initial estimate of the optimal solution x_0 and proceeds to find a better estimate of the solution each iteration. The search direction p_k at stage 'k' is given by the analogue of the solution of the Newton equation

$B_k P_k$ = -∇f(x_k)

Where B_k is the iterative approximation of the Hessian matrix of the problem and ∇f(x_k) is the gradient for the function evaluated at the position x_k. With respect to the two dimensional molecular optimisation problem, the gradient of the function at each point is the overall impact rotating the angle of the atom a_k with respect to the atom a_{k-1} on the total energy of the system. An alternative method for energy minimisation presented by

CASTEP is damped molecular dynamics, which involves only internal coordinates for each atom and uses critical damping to deal with the ground state. This can be implemented by using either a single damping coefficient for all degrees of freedom for coupled nodes or by using coefficients for each degree of freedom for independent nodes. However, for this problem it can be observed that BFGS algorithms are much more commonly utilised than gradient descent methods such as stochastic gradient descent as they generally require fewer iterations to reach an acceptable solution state.

Crossovers are a particularly significant issue which can arise during the run-time of a solving algorithm, and involves the crossing over of two different line segments in a bond constrained model. Crossovers are difficult to detect in a way which is not computationally expensive, and can often result in a state becoming close to but unable to reach the optimal system energy. They are particularly difficult to avoid in population based search implementations, as they rely on having a large pool of solution states which are combined at every iteration and presere the environments closest to the desired solution at each iteration. Due to the nature of these algorithms involving random combinations of atom positions, it is extremely likely for these combinations to introduce crossovers. Crossovers can be fixed by recursively straightening out the angle to the next atom in the chain associated with the offending atoms until no more crossovers occur, however it is possible for this to involve almost completely resetting a solution state to its starting positions. In general, it is better to avoid causing crossovers in the first place.

This review's purpose is to summarise existing research into molecular structure optimisation, as well as define key words and algorithms which are frequently used with respect to this problem and associated concepts. There are few studies which have been performed on the two dimensional molecular optimisation problem as it is effectively just a simplication of the three dimensional molecular optimisation problem, which is far more useful for providing models which can be used in real experiments. However, the algorithms which can be used to solve the three dimensional problem which have been discussed can be applied in a similar fashion to the two dimensional problem.

3 Algorithm Description

4 Results Discussion

5 Bibliography

Dictionary, a. (2018). atom Meaning in the Cambridge English Dictionary. [online] Dictionary.cambridge.org. Available at: https://dictionary.cambridge.org/dictionary/english/atom [Accessed 24 Jul. 2018]

Dictionary, m. (2018). molecule Meaning in the Cambridge English Dictionary. [online] Dictionary.cambridge.org. Available at: https://dictionary.cambridge.org/dictionary/english/molecule [Accessed 24 Jul. 2018]

shodor.org. (2018). Background Reading for Geometry Optimizations. [online] Available at: https://www.shodor.org/chemviz/optimization/students/background.html [Accessed 24 Jul. 2018]

Structure.usc.edu. (2018). Minimization and Molecular Dynamics. [online] Available at: http://structure.usc.edu/mmtk/MMTK_4.html [Accessed 24 Jul. 2018]

Jean, M. (2018). Energy Minimization Methods. [online] About.illinoisstate.edu. Available at: https://about.illinoisstate.edu/standard/Documents/CHE%20380.37/Handouts/380.37emin.pdf [Accessed 25 Jul. 2018].

Clark, S. (2018). Geometry Optimization. [online] Cmt.dur.ac.uk. Available at: http://cmt.dur.ac.uk/sjc/thesis_dlc/node37.html [Accessed 26 Jul. 2018].

Ul-Haq, Z. (2018). Introduction to Geometry Optimization. [online] Th.fhi-berlin.mpg.de. Available at: https://th.fhi-berlin.mpg.de/sitesub/meetings/dft-workshop-2016/uploads/Meeting/May_6_Qasmi.pdf [Accessed 26 Jul. 2018].

Spindynamics.org. (2018). molecular geometry optimization. [online] Available at: http://spindynamics.org/documents/cqc_lecture_6.pdf [Accessed 26 Jul. 2018].

Helgaker, T. (2009). Geometry optimization. [online] Folk.uio.no.
Available at: http://folk.uio.no/helgaker/talks/ESQC09_Optimization.pdf [Accessed 26 Jul. 2018].

Openmopac.net. (2018). The BFGS function optimizer. [online]
Available at: http://openmopac.net/manual/BFGS_optimizer.html [Accessed 26 Jul. 2018].

Tcm.phy.cam.ac.uk. (2018). CASTEP Geometry optimization. [online]
Available at: https://www.tcm.phy.cam.ac.uk/castep/documentation/WebHelp/content/modules/castep/thcastepgeomopt.htm
[Accessed 26 Jul. 2018].

Cpmd.org. (2018). Geometry Optimization. [online]
Available at: http://www.cpmd.org:81/manual/node80.html [Accessed 26 Jul. 2018].

General methods for geometry and wave function optimization
Thomas H. Fischer and Jan Almlof
The Journal of Physical Chemistry 1992 96 (24), 9768-9774
DOI: 10.1021/j100203a036

Tcm.phy.cam.ac.uk. (2018). CASTEP. [online]
Available at: https://www.tcm.phy.cam.ac.uk/castep/documentation/WebHelp/content/modules/castep/abtcastep.htm
[Accessed 26 Jul. 2018].

Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. J.; Refson, K.; Payne, M. C.
"First principles methods using CASTEP", Zeitschrift fuer Kristallographie, 220 (5-6), 567-570 (2005).