Computational Knowledge meets Quantum Chemistry

Stefan Janecek
uni software plus GmbH
Density Functional Theory

Computational Knowledge

Wolfram Language™
Chemistry

Drug Design

Molecular Biology

Material Science
(Steel production down to nanotechnology)

Semiconductor technology

\[ H\psi = E\psi \]
Ethane \((\text{C}_2\text{H}_6)\)

\[\psi(r_1, \ldots, r_N)\]

30 electrons

\[\psi(x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_{30}, y_{30}, z_{30})\]

20 grid points per coordinate

needs 8 \(20^{90}\) bytes
Ultimate storage medium (hypothetical): 1 byte per atom…

Volume to store $8 \times 20^{90}$ bytes?

$10^9$ Petabyte!
About $1.6 \times 10^{39} \text{ ly}^3$
(yes, cubic light-years)

Roughly $\sim 1 \text{ Million } x$
size of the universe
Density Functional Theory
(“DFT”)

Approximation to the *ground state* of the system

Based on two theorems by Pierre C. Hohenberg and Walther Kohn (1964)

Pierre C. Hohenberg
Walther Kohn (nobel laureate chemistry 1998)
John A. Pople (nobel laureate chemistry 1998)
Hohenberg-Kohn 1: “The density is special”

Whole information on ground state is “in principle” contained in the density alone!

\[ \psi(r_1, \ldots, r_N) \]

(100 kB … 100 MB)

(existence theorem alert !)

(Millions of universes….)
Everything depends on the density *alone*!

“Everything is a *functional* of the density”

\[ E = E[\rho] \]
Hohenberg-Kohn 2: “The energy is special”

The ground state density minimises the energy functional

\[ E[\rho \neq \rho_0] \geq E[\rho_0] = E_0 \]

Existence theorem alert:
we still don’t know the exact energy functional

Minimize approximations to the energy functional
(50 years of experience in finding good approximations)
Calculations: Demanding, but doable

<table>
<thead>
<tr>
<th></th>
<th>CPU time (core hours)</th>
<th>Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethane (30 electrons)</td>
<td>seconds</td>
<td>few megabytes</td>
</tr>
<tr>
<td>Fullerene (C60) (240 electrons)</td>
<td>5 core-h</td>
<td>few gigabytes</td>
</tr>
<tr>
<td>Si Nanowire (~57000 electrons)</td>
<td>1.1 Mio. core-h</td>
<td>few TB (?)</td>
</tr>
</tbody>
</table>
Show me what you can do!
Example 1: Quantum Dots
Example 2: Nuclear Magnetic Resonance
Benzene Molecule

Calculation Input:
• 6 Carbon
• 6 Hydrogen
• rough structure
chem. structure

NMR spectrum
DFT and the Wolfram Language
Why Wolfram Language

- Traditional languages: Fortran, C++
  - How do they compare?
- Rapid Development, framework that makes it easy to try out new ideas
- Use it as a “computational knowledge tool” (more later …)
Disclaimer:

This is a very rough, qualitative, anecdotal, largely inadequate comparison. It’s a “user story”, not a benchmark.

Only orders of magnitude count!
(i.e. a factor of 2 is not a big deal - a factor of 100 might be significant)
Efficiency: Runtime

Disclaimer: Very different Algorithms!
(only orders of magnitude count)

Mathematica code roughly $\sim$ C++ code

<table>
<thead>
<tr>
<th></th>
<th>C++ Code</th>
<th>Mathematica</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harmonic Oscillator</td>
<td>4.1s</td>
<td>4.5s</td>
</tr>
<tr>
<td>Fullerene (C60)</td>
<td>5h 35m</td>
<td>5h 55m</td>
</tr>
<tr>
<td>Full calculation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>SLOC</td>
<td>person years</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>--------------</td>
</tr>
<tr>
<td>LiMeReC</td>
<td>Fortran</td>
<td>50,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEST</td>
<td>C++</td>
<td>20,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematica DFT</td>
<td>Wolfram</td>
<td>300 (DFT) + 860 (FEM)</td>
</tr>
</tbody>
</table>
DFT and Computable Knowledge

- Chemical Space (“all possible compounds that can be produced by elements from the periodic table”) is huge
- No way to get empirical (experimental) data on all that
- At least a subset of that can be made **computable** (from just the knowledge of the chemical structure)
Vision: Quantum Chemistry and Computational Knowledge
Thank You!

Questions?