

The Hacker Within Group Project Proposal: Solid Hydrogen Atomic Trajectory

Yubo “Paul” Yang¹

¹*Department of Physics, University of Illinois, Urbana, Illinois 61801, USA*

(Dated: April 9, 2020)

I. INTRODUCTION

Properties and phase transitions of hydrogen under megabar pressures is important in diverse fields of study. For astronomy, models of the interior of giant gas planets such as Jupiter and Saturn depend critically on the nature of the liquid-liquid transition (LLT), namely whether it is first-order or continuous^{1,2}. For condensed matter, metallic hydrogen holds promise for a room temperature conventional superconductor^{3,4}. For computational physics, element number one remains an important benchmark for both electronic structure⁵ and ion dynamics methods.

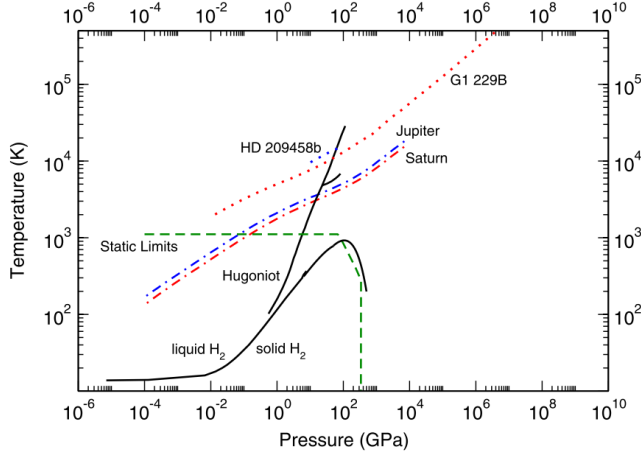


FIG. 1. Pressure-Temperature phase diagram of hydrogen from Ref.⁴.

Established experimental results on high-pressure hydrogen are limited in variety. At room temperature and below, diamond anvil cell (DAC) is the dominant apparatus to achieve such high pressures. Small size of the cell and fragility of the sample limit experimental probes to low-power optics such as infrared and Raman spectroscopy⁶. Hydrogen is a weak scatterer of X-Rays⁷, thus excluding this excellent tool for structural determination in most experiments. Only recently has X-ray been performed up to 254 GPa^{8,9}. At high temperatures, shock wave compression is the main method to achieve megabar pressures. Due to the transient nature of these experiments, acquiring and analyzing shock-wave data is challenging. Most notably, one cannot directly measure temperature, which may cause misinterpretation of raw data¹⁰⁻¹². Given the experimental difficulties, predictive simulations are highly desirable as they can inform and verify experiments¹³.

Simulation of high-pressure hydrogen is also challenging. Without experimental structural information from X-ray, many theoretical calculations have been performed on structures found in density functional theory (DFT) random structure searches¹⁴. Constraint by computational cost, these searches are limited to Classical protons, causing the methods to miss, for example, saddle-point structures that can be stabilized by nuclear quantum effect¹⁵. Most theoretical studies of high-pressure hydrogen suffer a similar dilemma between accuracy and practicality. On the one hand, predictive simulations of hydrogen require accurate methods both in the description of the electronic ground-state Born-Oppenheimer (BO) potential energy surface (PES) and in the inclusion of nuclear quantum effect beyond the quasi-harmonic approximation. The popular Perdew-Burke-Ernzerhof (PBE) density functional (DF) in DFT erroneously predicts some molecular structures to be metallic¹⁶. However, its use in conjunction with Classical molecular dynamics (MD) results in reasonable transition pressure for the LLT due to error cancellation¹⁷. This and other fortuitous cancellations of error has lead many to believe that the PBE functional provides a good description of solid hydrogen and caused much confusion in the community. PBE predicts a conductive molecular structure above 200 GPa, a molecular-to-atomic transition around 300 GPa¹⁸, and low-temperature superconducting liquid. All these predictions contradict experimental evidence. Systematic benchmark of the PES from various DFT functionals against QMC found the vdW-DF1 functional to be the most accurate for molecular hydrogen¹⁹. However, this functional has yet to gain widespread adoption due to higher computational cost and lower popularity compared to PBE.

In summary, the properties of hydrogen over a large pressure and temperature range are interesting yet challenging to characterize in both theory and experiment. On the other hand, to explore such large pressure and temperature ranges using first-principle methods is not only time consuming, but also prone to other limitations such as inadequate treatment of finite-size effect, nuclear quantum effect, and simulation timescale. Further, some experimentally measurable properties, such as emission, absorption, infrared, and Raman spectra, heat conductivity, and diffusion constants are difficult to calculate in the most accurate electronic methods, e.g., DMC. In recent years, machine learning has emerged as a promising tool for striking a balance between accuracy and efficiency in these studies. One can learn the PES of first-principle calculations, then use MD to simulate large systems and calculate properties efficiently.

However, programs for constructing and validating these machine-learned potentials are still in their early stages. As such, tools for visualizing the accuracy of these models is invaluable for the community.

II. DATA DESCRIPTION

The bulk of the database will consist of atomic snapshots sampled during first-principle MD simulations. These data are publicly available in the “melth2” folder on yt Hub. They can be retrieved in scripts via the RESTful APIs available for yt Hub. Most data are prepared using the popular atomic simulation environment (**ase**) Python package, so atomic snapshots can be easily imported and exported using the **ase.io** submodule. Details of the database are described in the “README.md” file in the above folder and will be updated as needed. In the following, I reproduce parts of the “README.md” file to make this proposal a self-contained document.

An atomic configuration is defined by a list of atomic species, each associated with a spatial coordinate. We typically store configurations in QUIP-style extended xyz format, for example:

```
2
h2 molecule
H 0.0 0.0 0.0
H 0.0 0.0 0.74
```

defines a hydrogen molecule in vacuum, whereas

```
2
Lattice='5.0 5.0 5.0'\
Properties=species:S:1:pos:R:3\
pbc='T T T'\
H 0.0 0.0 0.0
H 0.0 0.0 0.74
```

defines a hydrogen molecule in a cubic box with periodic boundary conditions. An xyz file always starts with the number of atoms, followed by a description line, then a list of all the atoms. Not all programs parse the description line of an xyz file, so we sometimes have to separately parse box and boundary conditions. Units are angstrom and eV for length and energy, respectively.

Given an atomic configuration, one can evaluate

- total energy
- forces on each atom
- stress tensor on the simulation cell

The virial tensor is $\text{stress} \times \text{volume}$.

These properties can be evaluated using various electronic structure methods and serve as input data to PES fitting algorithms in the next step.

We typically store energy, forces, and stress in QUIP-style extended xyz format, for example:

```
96
Lattice='5.4463 6.2889 5.755'\
Properties=species:S:1:pos:R:3:forces:R:3\
energy=-1451.9548200705292\
stress='-0.91683 0.05019 0.00608\
0.05019 -0.91530 -0.00435\
0.00608 -0.00435 -0.925558'\
pbc='T T T'\
H -0.82266035 0.48413997 0.69739698 2.19773810
H 1.00888387 2.05069199 0.91559926 -0.68396464
...
```

, which can be written by **ase**. Here, we concatenate these xyz files and convert to binary traj files before upload. We use traj instead of CDF format because **ase** does not support writing forces and stress to CDF format. Units are angstrom and eV for length and energy, respectively.

III. GROUP PROJECT PROPOSAL

A. Front-end

The front-end development will mostly involve communication between the data source (yt Hub), the visualization engine (Python?), and the web interface (React?). This code will also allow us to access the many astronomy data collections already available on yt Hub for future projects.

B. Visual Representation

3D visualization is the bread and butter of atomic trajectory representation. The most basic functionality of our visualization app. should be to display an atomic snapshot at a user-specified thermodynamic condition (temperature and pressure). Further functionalities can include:

1. show movie of atom motion along the trajectory
2. show force error of ML PES by coloring the atoms
3. show average particle density as volumetric effect

C. Audio Representation

One idea to add audio to the representation of atomic trajectory is to assign harmonizing frequencies to the diagonal of the stress tensor and allow change in stress to shift the initial frequencies. This way, a simulation with isotropic stress will sound harmonious, whereas one with anisotropic stress will sound “off”, prompting the researcher to check data.

The same strategy can be adopted to check the accuracy of the prediction of stress tensor by ML PES, i.e.

make the simulation sound harmonious when the error is low, but “off” when the error is high.

D. “Physical” Representation

While a physical model of an atomic snapshot is unlikely to be informative, one can import the model into a

Game engine for interactive virtual reality (VR) or augmented reality (AR) experiences. One can explore using haptic feedback to simulation a “physical” model.

-
- ¹ W. B. Hubbard and B. Militzer, *Astrophys. J.* **820**, 80 (2016), arXiv:1602.05143.
 - ² S. M. Wahl, W. B. Hubbard, B. Militzer, T. Guillot, Y. Miguel, N. Movshovitz, Y. Kaspi, R. Helled, D. Reese, E. Galanti, S. Levin, J. E. Connerney, and S. J. Bolton, *Geophys. Res. Lett.* **44**, 4649 (2017), arXiv:1707.01997.
 - ³ J. M. McMahon and D. M. Ceperley, *Phys. Rev. B* **84**, 144515 (2011), arXiv:1106.5526.
 - ⁴ J. M. McMahon, M. A. Morales, C. Pierleoni, and D. M. Ceperley, *Rev. Mod. Phys.* **84** (2012), 10.1103/RevModPhys.84.1607.
 - ⁵ M. Motta, D. M. Ceperley, G. K.-L. Chan, J. A. Gomez, E. Gull, S. Guo, C. A. Jiménez-Hoyos, T. N. Lan, J. Li, F. Ma, A. J. Millis, N. V. Prokof'ev, U. Ray, G. E. Scuseria, S. Sorella, E. M. Stoudenmire, Q. Sun, I. S. Tupitsyn, S. R. White, D. Zgid, and S. Zhang, *Phys. Rev. X* **7**, 031059 (2017), arXiv:1705.01608.
 - ⁶ R. P. Dias and I. F. Silvera, *Science* (80-.). **355**, 715 (2017), arXiv:1702.04246.
 - ⁷ C.-S. Zha, R. E. Cohen, H.-K. Mao, and R. J. Hemley, *Proc. Natl. Acad. Sci. U. S. A.* **69**, 1129 (2014).
 - ⁸ Y. Akahama, M. Nishimura, H. Kawamura, N. Hirao, Y. Ohishi, and K. Takemura, *Phys. Rev. B* **82**, 060101 (2010).
 - ⁹ C. Ji, B. Li, W. Liu, J. S. Smith, A. Majumdar, W. Luo, R. Ahuja, J. Shu, J. Wang, S. Sinogeikin, Y. Meng, V. B. Prakapenka, E. Greenberg, R. Xu, X. Huang, W. Yang, G. Shen, W. L. Mao, and H.-K. Mao, *Nature* **573**, 558 (2019).
 - ¹⁰ P. M. Celliers, M. Millot, S. Brygoo, R. S. McWilliams, D. E. Fratanduono, J. R. Rygg, A. F. Goncharov, P. Loubeyre, J. H. Eggert, J. L. Peterson, N. B. Meezan, S. Le Pape, G. W. Collins, R. Jeanloz, and R. J. Hemley, *Science* (80-.). **361**, 677 (2018).
 - ¹¹ M. D. Knudson, D. L. Hanson, J. E. Bailey, C. A. Hall, J. R. Asay, and C. Deeney, *Phys. Rev. B* **69**, 144209 (2004).
 - ¹² M. D. Knudson and M. P. Desjarlais, *Phys. Rev. Lett.* **118**, 035501 (2017).
 - ¹³ C. Pierleoni, M. A. Morales, G. Rillo, M. Holzmann, and D. M. Ceperley, *Proc. Natl. Acad. Sci.* **113**, 4953 (2016).
 - ¹⁴ C. J. Pickard and R. J. Needs, *Nat. Phys.* **3**, 473 (2007).
 - ¹⁵ B. Monserrat, R. J. Needs, E. Gregoryanz, and C. J. Pickard, *Phys. Rev. B* **94**, 134101 (2016), arXiv:1609.07486.
 - ¹⁶ N. D. Drummond, B. Monserrat, J. H. Lloyd-Williams, P. L. Ríos, C. J. Pickard, and R. J. Needs, *Nat. Commun.* **6**, 7794 (2015).
 - ¹⁷ M. A. Morales, J. M. McMahon, C. Pierleoni, and D. M. Ceperley, *Phys. Rev. Lett.* **110**, 1 (2013), arXiv:arXiv:1303.6671v1.
 - ¹⁸ J. McMinis, R. C. Clay, D. Lee, and M. A. Morales, *Phys. Rev. Lett.* **114**, 1 (2015).
 - ¹⁹ R. C. Clay, M. Holzmann, D. M. Ceperley, and M. A. Morales, *Phys. Rev. B* **93**, 035121 (2016).