

Group Meeting - December 18, 2020

Paper review & Research progress

Truong Son Hy *

*Department of Computer Science
The University of Chicago

Ryerson Physical Lab



Erwin Schrödinger

Consciousness cannot be accounted for in physical terms. For consciousness is absolutely fundamental. It cannot be accounted for in terms of anything else.

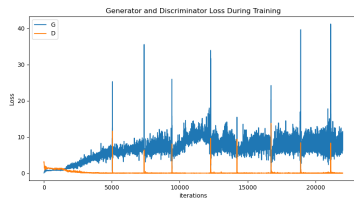
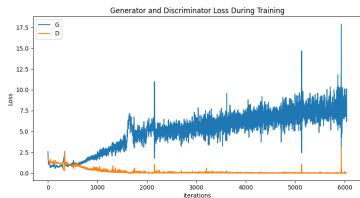


The Schrödinger's cat lives or dies with equal probability $(1/\sqrt{2})^2 = 1/2$.
Until we measure the cat, he/she exists in **superposition**.



Research update (1)

I am currently trying to use GAN for the task of molecular generation. But training for GAN (the mini-max game) is **tricky**!



QM9 (left), ZINC (right).
Generator (G), Discriminator (D).

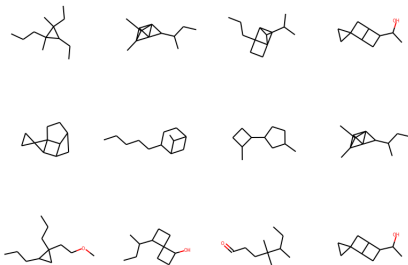
Problem: D does very well in distinguishing fake/non-fake examples, while G only generates crap molecules. We need some way to initialize Generator.



Research update (2)

Thus, I apply the **pre-training** mixing VAE and GAN:

- 1 Start by training a VAE (or MRF) with an encoder and decoder for few epochs.
- 2 Then, take the decoder of VAE to be inside the generator of GAN. And continue training GAN.



Result: Validity 81.9%, Unique 14.64% and 23 long molecules among 5000 generated.



- 1 **Differentiable Molecular Simulations for Learning and Control** (NeurIPS 2020 workshop), https://ml4molecules.github.io/papers2020/ML4Molecules_2020_paper_32.pdf
- 2 **The Schrödinger Equation** (Terrence Tao), <https://www.math.ucla.edu/~tao/preprints/schrodinger.pdf>



The Schrödinger Equation

Terrence Tao

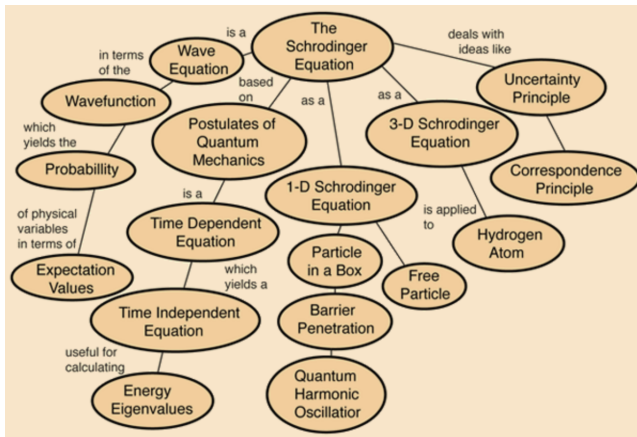
<https://www.math.ucla.edu/~tao/preprints/schrodinger.pdf>

Note

I realized that to understand the **Differentiable Molecular Simulations for Learning and Control** paper, I need to have an intuition for the Schrödinger equation, so I choose to follow the note from Terrence Tao.



The Schrödinger Equation – Tree of concepts



Source: <http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/schrcn.html>

<http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/schrcn.html>



The Schrödinger Equation – Classical mechanics (1)

Assumption

- Consider a **single** particle with some mass $m > 0$ in n -dimensional space \mathbb{R}^n moving subject to the influence of a potential (as a function) $V : \mathbb{R}^n \rightarrow \mathbb{R}$.
- Assume all functions are smooth (infinitely differentiable).

The **state** of a system at any given time t is described by the element $(q(t), p(t))$ of **phase space** $\mathbb{R}^n \times \mathbb{R}^n$ where:

- $q(t) \in \mathbb{R}^n$ is a specific position.
- $p(t) \in \mathbb{R}^n$ is a specific momentum.
- $p(t) = mv(t)$ and $v(t) = q'(t)$ where $v(t)$ is the velocity of the particle.



The Schrödinger Equation – Classical mechanics (2)

The **energy** of the state is described by the **Hamiltonian function** $H : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ on phase space:

$$H(q, p) = \frac{|p|^2}{2m} + V(q) = \text{kinetic energy} + \text{potential energy}$$

The system evolves by **Hamilton's equations of motion**:

$$q'(t) = \frac{\partial H}{\partial p}, \quad p'(t) = -\frac{\partial H}{\partial q}$$

From the **conservation of energy**:

$$\frac{d}{dt} H(q(t), p(t)) = 0$$

we see that there is a quantity E (independent of time) such that:

$$H(q(t), p(t)) = E \quad (\forall t \in \mathbb{R})$$



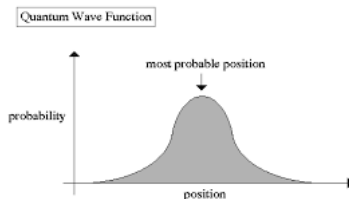
The Schrödinger Equation – Quantum mechanics (1)

In quantum mechanics, the state of the particle at a time t is no longer described by a single point $(q(t), p(t))$ in phase space, but described by a complex-valued **wave function** $\psi(t) : \mathbb{R}^n \rightarrow \mathbb{C}$ obeying the normalization

$$\langle \psi(t), \psi(t) \rangle = 1$$

where the inner product is defined as

$$\langle \phi, \psi \rangle = \int_{\mathbb{R}^n} \phi(q) \overline{\psi(q)} dq$$



$|\psi(t)|^2$ is a probability distribution of the particle's position



The Schrödinger Equation – Quantum mechanics (2)

The Probability that an electron lands somewhere between x_L and x_R is:

$$\text{Prob}(x_L \leq x \leq x_R) = \int_{x_L}^{x_R} P(x) dx = \int_{x_L}^{x_R} |\psi(x)|^2 dx$$

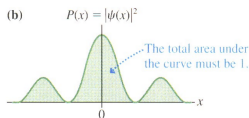
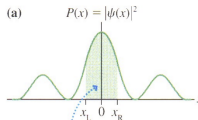


FIGURE 39.8 The area under the probability density curve is a probability.

Normalization

$$\int_{-\infty}^{\infty} P(x) dx = \int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1$$

Total Area must equal 1.

Richard Feynman, **Lectures on Physics** (volume III)

History

Erwin Schrödinger did not propose this interpretation of the wave function, but Max Born. The interpretation has been widely accepted since then.

The Schrödinger Equation – Quantum mechanics (3)

Unlike classical mechanics, a wave function $\psi(t)$ does not necessarily have a specific position $q(t)$, but it has an **average position** $\langle q(t) \rangle$ defined as:

$$\langle q(t) \rangle = \langle Q\psi(t), \psi(t) \rangle = \int_{\mathbb{R}^n} q |\psi(t, q)|^2 dq = \int_{\mathbb{R}^n} [Q\psi(t, q)] \overline{\psi(t, q)} dq$$

where the **position operator** Q is the operation of multiplication by q :

$$Q\psi(t, q) = q\psi(t, q)$$

Similarly, **average momentum**:

$$\langle p(t) \rangle = \langle P\psi(t), \psi(t) \rangle = \frac{\hbar}{i} \int_{\mathbb{R}^n} [\nabla_q \psi(t, q)] \overline{\psi(t, q)} dq$$

where the **momentum operator** P is defined by Planck's law:

$$P\psi(t, q) = \frac{\hbar}{i} \nabla_q \psi(t, q), \quad (\hbar = h/2\pi)$$



The Schrödinger Equation – Quantum mechanics (4)

Quantum observable

A **quantum observable** is understood as a self-adjoint operator A acting on the space of complex-valued square integrable functions, such that we can define the **average value** $\langle A(t) \rangle$:

$$\langle A(t) \rangle = \langle A\psi(t), \psi(t) \rangle$$

The analogue of Hamilton's equations of motion is the **time-dependent Schrödinger equation**:

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi$$

but H is now **quantum observable**:

$$H = \frac{|P|^2}{2m} + V(Q)$$



The Schrödinger Equation – Quantum mechanics (5)

Time-dependent Schrödinger equation:

$$i\hbar \frac{\partial \psi}{\partial t}(t, q) = H\psi(t, q) = -\frac{\hbar^2}{2m} \Delta_q \psi(t, q) + V(q)\psi(t, q)$$

where the Laplacian of ψ is defined as:

$$\Delta_q \psi = \sum_{j=1}^n \frac{\partial^2 \psi}{\partial q_j^2}$$

The conservation of energy law:

$$\frac{d}{dt} \langle H(t) \rangle = 0$$



The Schrödinger Equation – Quantum mechanics (6)

If there exists a real number E such that the quantum state ψ oscillates in time according to the formula:

$$\psi(t, q) = e^{\frac{E}{i\hbar}t} \psi(0, q)$$

we have:

$$H\psi(t) = E\psi(t) \quad (\forall t)$$

The equation:

$$H\psi = E\psi$$

is known as **time-independent Schrödinger equation** where E is referred to as the **energy level** or **eigenvalue** of the state ψ .



The Schrödinger Equation – Quantum mechanics (7)

In general, the quantum state ψ has instead a **super-position** or **linear combination** of oscillating states:

$$\psi(t, q) = \sum_{k=1}^K e^{\frac{E_k}{i\hbar} t} \psi_k(q)$$

for a finite number of energy levels E_1, \dots, E_K ; and a finite number of functions ψ_1, \dots, ψ_K each obeying the time-independent Schrödinger equation at energy level E_k : $H\psi_k = E_k\psi_k$.

Note

This looks like a Fourier transformation to me:

$$\psi(t, q) = \sum_{k=1}^K e^{-i\frac{E_k t}{\hbar}} \psi_k(q)$$

The Schrödinger Equation – Quantum mechanics (8)

Fourier transform

- Knowing all the solutions to the time-independent Schrödinger equation gives all the solutions to the time-dependent Schrödinger equation.
- Two equations can be viewed as Fourier transforms of each other. The quantity E/\hbar is the frequency variable dual to the time variable t .

Solving the time-independent equation

- The time-independent Schrödinger equation $H\psi = E\psi$ is an eigenvalue equation for the operator H .
- If H is a self-adjoint transformation on a finite-dimensional space, then there would only be a finite number of eigenvalues E for which the equation has a non-trivial solution.

The Schrödinger Equation – Quantum mechanics (9)

Example 1: The time-independent Schrödinger equation $H\psi = E\psi$. Consider a single wave – energy eigenfunction ψ_n with energy E_n . We have the expectation:

$$\langle H \rangle = \int_{-\infty}^{\infty} \psi_n^* H \psi_n dx = \int_{-\infty}^{\infty} \psi_n^* E_n \psi_n dx = E_n \int_{-\infty}^{\infty} |\psi_n|^2 dx = E_n$$

Thus: $\langle H \rangle = E_n$ and $\langle H^2 \rangle = E_n^2$. We have the variance:

$$\sigma_H^2 = \langle H^2 \rangle - \langle H \rangle^2 = 0$$

That means **no** spread in energy for ψ_n , and we always measure the energy E_n every time.



The Schrödinger Equation – Quantum mechanics (10)

Example 2: Continue from Example 1. Suppose we have a **mixture** of waves ψ_n (in superposition):

$$\psi_{\text{mixture}} = \sum_n c_n \psi_n$$

Then:

$$H\psi_{\text{mixture}} = \sum_n c_n H\psi_n = \sum_n c_n E_n \psi_n$$

If we measure the energy of ψ_{mixture} then: we will find energy E_n with **probability** $|c_n|^2$.



The Schrödinger Equation – Solution for Hydrogen atom

MIT OpenCourseWare video lecture:

<https://www.youtube.com/watch?v=KfbvrGt3MlI>

Other sources:

[http:](http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/hydsch.html)

[//hyperphysics.phy-astr.gsu.edu/hbase/quantum/hydsch.html](http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/hydsch.html)

<https://users.aber.ac.uk/ruw/teach/327/hatom.php>



Differentiable Molecular Simulations for Learning and Control (NeurIPS 2020 workshop)

Wujie Wang, Simon Axelrod, Rafael Gómez-Bombarelli

https://ml4molecules.github.io/papers2020/ML4Molecules_2020_paper_32.pdf

Note: I want to learn about the molecular simulation by presenting this work. Please correct me to improve my understanding!



Introduction

Non-quantized molecular behavior is typically simulated with differential equations parameterized by a Hamiltonian H , or energy function. The Hamiltonian describes the state of the system and its interactions with the environment.

The Hamiltonian of the systems (more than one particles)

Classically, simulating the positions of points that preserves energy requires integrating the Hamiltonian equations of motions:

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}$$

$$H(\mathbf{p}, \mathbf{q}) = U(\mathbf{q}) + \sum_i^N \frac{p_i^2}{2m_i}$$

where p_i , q_i and $p_i^2/(2m_i)$ are the respective momentum, position and kinetic energy of the i -th particle. $U(\mathbf{q})$ is the potential energy.

Proposals

Proposals

- **Context:** Use differentiable simulations in the context of molecular simulation.
- **Claim:** A Hamiltonian can be learned such that the macroscopic observables computed through simulation trajectory to match a given target.
- **Method:** Automatic differentiation of the macroscopic observables computed from simulation trajectory with respect to a wide class of functional forms for the energy function including Graph Neural Networks (GNN). → The Hamiltonian is modeled by a parameterized GNN.



Overview (1)

The total Hamiltonian of the system and the control field is given by:

$$\hat{H}(t) = \hat{H}_S + \hat{H}_b$$

where \hat{H}_S is the system Hamiltonian, and \hat{H}_b is the control Hamiltonian. The system wave function evolves under the Schrödinger equation:

$$\frac{\partial}{\partial t}|\psi(t)\rangle = -i\hat{H}(t)|\psi(t)\rangle$$

where $|\psi(t)\rangle$ is the wave function.

- The authors proposed to use a GNN to represent the control Hamiltonian H_b .
- They back-propagate through the simulations to continuously update the GNN such that the loss function (measuring bond distances, angles and dihedrals angles) is minimized.



Overview (2)

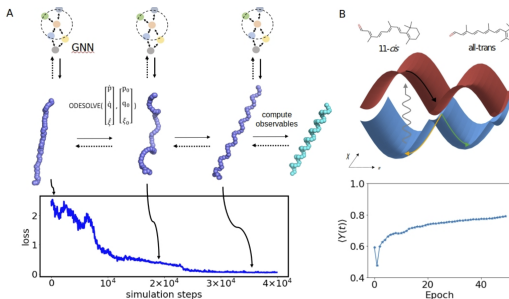


Figure 1: **A** We perform continuous model training during the simulations to bias a harmonic polymer chain toward a targeted helix shape. This is done by training a bias Hamiltonian parameterized by a GNN. We run the simulations for 4000 steps, and the loss is computed and differentiated to update GNN weights every 40 simulation steps. **B** Controlled isomerization of the model retinal Hamiltonian with a time-dependent electric field. The model consists of two electronic states, denoted with blue and red, a vibrational mode x , and a torsional mode ϕ . We show that the time-averaged quantum yield increases as a function of training epoch with differentiable control.



Learning from observables

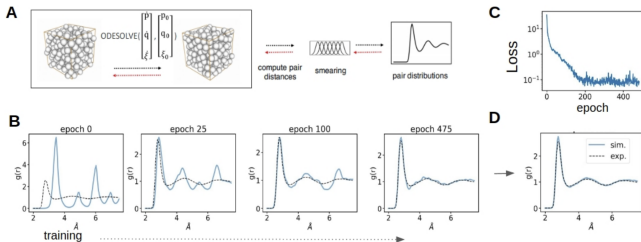
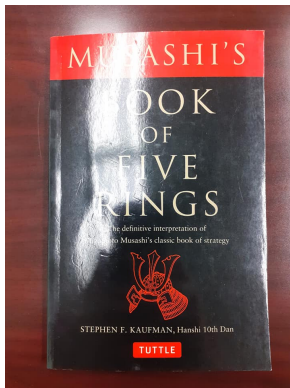


Figure 2: A. Computational workflow to fit Oxygen-Oxygen pair distribution functions for water at 298K and 1 atm. We use a Graph Neural Networks to parameterize the force fields. For each training epoch, we compute the pair distribution functions from simulated trajectories. We back-propagate the mean square loss between simulated and target pair distribution functions to update the GNN parameters. B. Change of simulated g_{oo} during training. C. The training loss curve. D. g_{oo} simulated by the converged model.



The Book of the Year 2020

- 1 The ultimate aim of martial arts is **not** having to use them.
- 2 You must understand that there is more than one path to the top of the mountain.
- 3 Think lightly of yourself and deeply of the world.



The Book of Five Rings – Miyamoto Musashi

