OUR Journal: ODU Undergraduate Research Journal

Volume 11 Article 3

October 2024

An Introduction to the Time-Independent Schrödinger Equation and Methods to Solve it

Vu Giang
Old Dominion University

Alex Gnech
Old Dominion University

Recommended Citation

Giang, Vu and Gnech, Alex (2024) "An Introduction to the Time-Independent Schrödinger Equation and Methods to Solve it," *OUR Journal: ODU Undergraduate Research Journal*: Vol. 11, Article 3.
DOI: https://doi.org/10.25778/9dd7-vn32
Available at: https://digitalcommons.odu.edu/ourj/vol11/iss1/3

An Introduction to the Time-Independent Schrödinger Equation and Methods to Solve it

Cover Page Footnote

I thank my advisor Dr. Gnech and my family for supporting me along this journey and encouraging me to pursue my academic endeavors.

OLD DOMINION UNIVERSITY UNDERGRADUATE RESEARCH JOURNAL VOL. 11, OCTOBER 2024

An Introduction to the Time-Independent Schrödinger Equation and Methods to Solve it

Vu Giang, The Department of Mathematics and Statistics, Old Dominion University
Alex Gnech, The Department of Physics, Old Dominion University

***Abstract*—The Time-Independent Schrödinger Equation is a linear elliptic PDE that describes quantum-mechanical systems. Its significance in the science of submicroscopic phenomena, particularly quantum mechanics, is as central as Newton's laws of motion are to classical mechanics. This study uses various methods, including novel neural networks and finite difference schemes, to solve the one-dimensional two-body equation.**

***Index Terms*—deuteron, harmonic oscillator, ladder operator, mathematical physics, Metropolis-Hastings, neural network, numerical methods, numerov, optimization, partial differential equation, Schrödinger equation, finite difference**

I. Introduction

WHEN one hears about quantum mechanics, whether the layperson or a physics undergraduate student, the typical consensus is this field of study challenges many and is out of reach for others [4], [12]. Hopefully, this article can give a more intuitive understanding drawn from historical and mathematical backgrounds. One article, [12] criticizes many textbooks and professors introducing this subject unintuitively and with no motivation for the students to hold on to it. Relating quantum mechanics to something the student may know of will allow for a quicker and deeper grasp of the subject.

Generally, quantum mechanics describes with good approximation microscopic phenomena. No theory or field of study has been able to succinctly describe or predict atomic behavior comparable to quantum mechanics. Some examples include quantum mechanics explaining why pigments of leaves are the way they are, explaining the rigidity of certain plastics or metals, or making medical devices a possibility. The motivation to do research in this field is to further the understanding of natural phenomena and to develop innovative technology.

To gain a better understanding, the wave is introduced with a look into history. Quantum physics' creation is attributed to German physicist Max Planck, who suggested that the energy emitted from a black body is quantized. Essentially, energy is associated with the emitted frequency through the equation $E = hf = \frac{hc}{\lambda}$, where E is the energy, h is the Planck constant, f is the frequency of the oscillation, c is the speed of light, and λ is the wavelength of the light. Then came Albert Einstein, who wondered why electrons radiate energy in quanta. He deduced that radiation was quantized into particles called photons, each of which has a limit to carry certain quanta of energy, explaining the results of the photoelectric effect observed by Heinrich Hertz and Wilhelm Hallwachs. When a constant beam of light hits a surface, the atoms in the affected region are excited by the photons hitting it by this transfer of energy from the collision of the photon to the atom.

However, increasing the intensity of the beam does not change the energy of the emitted photon, proving that electrons in atoms have quantized energy. Later, Arthur Compton performed an experiment outputting x-ray scattering off electrons, which changes the wavelengths. Tying this phenomenon with the Planck relation, he showed the particle nature of light (i.e. photons). Louis de Broglie furthered this idea beyond radiation to matter. He proposed that wave-particle duality is a property of nature described mathematically by $p = h/\lambda$ where p is the particle momentum and λ is the associated wavelength. The electron double-slit experiment results showed the scattering of waves of the electron proving de Broglie's claim.

To start talking about quantum mechanics, it is worthwhile to consider the waves produced by a pebble dropped into some water, as depicted in Fig. (1) and (2).

Instead of looking at the damped wave, consider the simplest wave:

$$S(r) = S_0 \frac{\sin(r)}{r}$$

$S(r)$ represents the displacement of the wave, while S_0 is the amplitude of the wave and r is the position of the

Manuscript submitted June 17, 2024, accepted September 17, 2024, revised October 1, 2024.

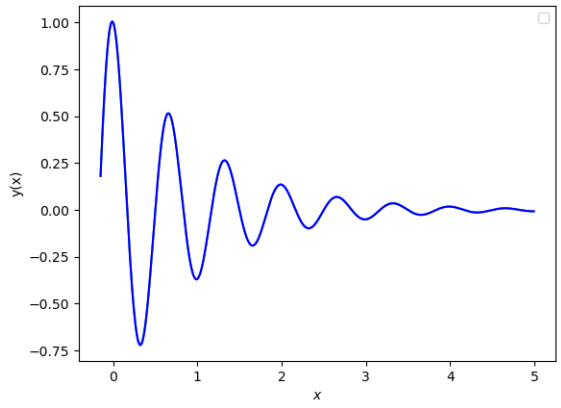


Fig. 1: A 2D depiction of a wave $y = e^{-x}\cos(3\pi x)$.

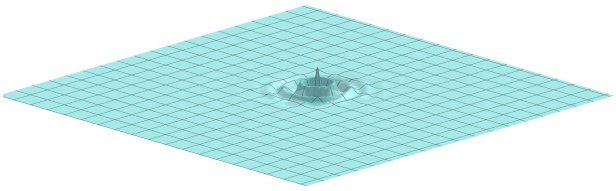


Fig. 2: A 3D depiction of a wave $z = e^{-\sqrt{2x^2+2y^2}}\cos(3\pi\sqrt{2x^2+2y^2})$.

function. The function can be rewritten to be even more general:

$$S(r,t) = S_0 \sin(kr - \omega t)$$

where k is the wave number and ω is the angular frequency. Note that the phase velocity of the wave is defined as $v = \omega/k$. Taking the second derivative of the equation above concerning r and t gives exactly the wave equation.

$$\frac{\partial^2 S(r,t)}{\partial t^2} = v^2 \nabla^2 S(r,t) \tag{1}$$

Also note that the cosine function describes a wave shifted by some phase, so generalizing the above result further, some complex variables are used to rewrite $S(r,t)$ as:

$$\begin{aligned} S(r,t) &= S_0 e^{i(kr-\omega t)} \\ &= S_0[\cos(kr-\omega t) + i\sin(kr-\omega t)] \end{aligned} \tag{2}$$

Recall that in quantum mechanics $p = h/\lambda$, $E = hf = \hbar\omega$, and knowing that $k = 2\pi/\lambda$ we can rewrite the wave number as $k = 2\pi p/h = p/\hbar$. Substituting it into (2) gives:

$$\Psi(r,t) = \Psi_0 e^{i(pr-Et)/\hbar}$$

To match how the wave function is normally represented, change S to Ψ. Understanding that the subject of this explanation is a non-relativistic free particle under a plane wave, the total energy of a particle is given by the kinetic energy $E = p^2/2m$ where m is the particle or system mass. Using this fact, the wave equation (1) now becomes

$$i\hbar\frac{\partial\Psi(r,t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(r,t)\,. \tag{3}$$

Keep in mind that only the first derivative with respect to time appears on the right-hand side. If (3) is further generalized, where the particle is under the influence of some potential. Then (3) can be interpreted as $E = T + V$, where T is the kinetic energy and V is the potential energy. In (3), the right-hand side is the kinetic energy, so adding the potential energy gives:

$$i\hbar\frac{\partial\Psi(r,t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(r,t) + V(r)\Psi(r,t) \tag{4}$$

Thus, producing the general Schrödinger equation. Similar derivations can be found in the following literatures Refs. [3], [5], [7], [12], [17]. Other considerations are important for the interpretation of quantum mechanics, but for this paper, we will only remark on the meaning of the wave function. We assume the Born Interpretation of Quantum Mechanics. Practically, the wave function is a density probability, meaning if there is a probability of finding a particle within some interval the probability is given by

$$P_L = \int_\Omega |\Psi(r)|^2 dr = 1\,.$$

For our discussion, this will be enough to begin the discussion of solving the Schrödinger equation.

II. Exact Solution

Before solving the Schrödinger equation using the numerical and neural network methods, we must establish an exact solution to compare against. For simplicity, we use the solution to the quantum harmonic oscillator. Numerous ways exist to solve this problem, and the chosen method uses the so-called ladder operators. Given the Schrödinger equation (4), reducing the equation to the 1-D case gives:

$$i\hbar\frac{\partial\Psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(x,t) + V(x)\Psi(x,t) \tag{5}$$

From Eq. (4), it's clear to see that the function Ψ is dependent on time. For this study ($\psi = e^{-iEt/\hbar}\tilde{\psi}$ this is how you pass from Eq. (4) to Eq. (7)), only the time-independent type equation is considered. Therefore, Ψ will now only depend on x, and the derivative term

https://digitalcommons.odu.edu/ourj/vol11/iss1/3
DOI: https://doi.org/10.25778/9dd7-vn32

with respect to time is now a constant. In this field, this derivative is seen as the energy, which we will see later. Thus, this gives:

$$i\hbar\frac{d\Psi(x)}{dt} = -\frac{\hbar^2}{2m}\frac{d^2\Psi(x)}{dx^2} + V(x)\Psi(x) \tag{6}$$

Now apply our potential, $V(x) = \frac{1}{2}m\omega^2x^2$ gives:

$$i\hbar\frac{d\Psi(x)}{dt} = -\frac{\hbar^2}{2m}\frac{d^2\Psi(x)}{dx^2} + \frac{1}{2}m\omega^2x^2\Psi(x) \tag{7}$$

The ladder operator method was inspired by Griffith [6]. If we see the left-hand side as the energy then we can rewrite the equation as follows:

$$E\Psi(x) = \frac{1}{2m}\left[(\frac{\hbar}{i}\frac{d}{dx})^2 + (m\omega x)^2\right]\Psi(x)$$

Define $a_\pm : \frac{1}{\sqrt{2m}}(\frac{\hbar}{i}\frac{d}{dx} \pm im\omega x)$. Since a is an operator, it is not commutative. Then it follows that

$$\begin{aligned}
a_-a_+ &= \frac{1}{2m}\left[(\frac{\hbar}{i}\frac{d}{dx})^2 + (m\omega x)^2\right] + \frac{1}{2}\hbar\omega \\
a_+a_- &= \frac{1}{2m}\left[(\frac{\hbar}{i}\frac{d}{dx})^2 + (m\omega x)^2\right] - \frac{1}{2}\hbar\omega \\
&\implies (a_-a_+ - \frac{1}{2}\hbar\omega)\Psi(x) = E\Psi(x) \\
&\implies (a_+a_- + \frac{1}{2}\hbar\omega)\Psi(x) = E\Psi(x) \\
&\implies a_-a_+ - a_+a_- = \hbar\omega
\end{aligned}$$

A proof of $a_\pm\Psi(x)$ satisfying the Schrödinger equation with energy $E \pm \hbar\omega$ is omitted but can be found in [6]. So, now we have $a_\pm\Psi(x) = E \pm \hbar\omega$. What this means is that each time the operator a acts on Ψ we get a new energy level, and since we can never have negative energy, there must be a lowest energy level meaning $a_-\Psi_0(x) = 0$, which is the ground state. Thus,

$$\begin{aligned}
a_-\Psi(x) &= \left[\frac{1}{\sqrt{2m}}\left(\frac{\hbar}{i}\frac{d}{dx} - im\omega x\right)\right]\Psi(x) \\
&= 0 \\
\implies &= \frac{1}{\sqrt{2m}}\left(\frac{\hbar}{i}\frac{d\Psi(x)}{dx} - i\Psi(x)m\omega x\right) \\
\implies \frac{d\Psi(x)}{dx} &= -\frac{m\omega x}{\hbar}\Psi(x)
\end{aligned}$$

With this equation, we can integrate both sides with respect to x.

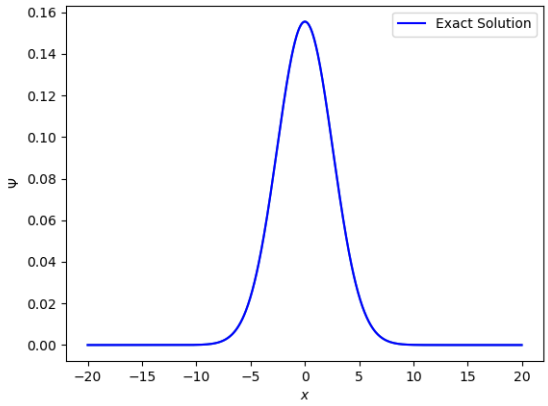


Fig. 3: The groundstate wave function Ψ_0 with m and $\omega = 1$.

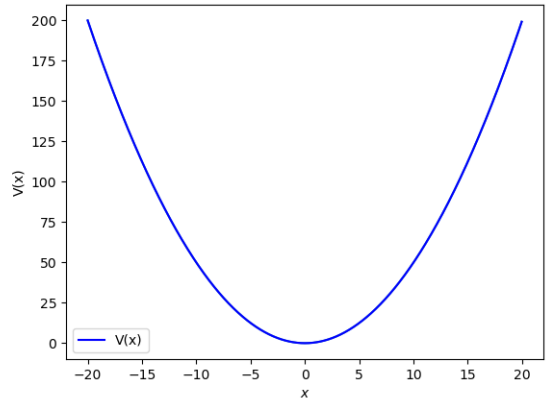


Fig. 4: The respective potential for Ψ_0, $V(x) = \frac{1}{2}m\omega x^2$.

$$\begin{aligned}
\int\frac{d\Psi(x)}{dx}dx &= -\int\frac{m\omega x}{\hbar}\Psi(x)dx \\
\implies \int\frac{d\Psi(x)}{\Psi(x)} &= -\frac{m\omega}{\hbar}\int x dx \\
\implies \ln|\Psi(x)| + C_1 &= -\frac{m\omega}{\hbar}\frac{x^2}{2} + C_2 \\
\implies \Psi(x) &= Ae^{-\frac{m\omega}{2\hbar}x^2}
\end{aligned}$$

Therefore, $\Psi_0(x) = A_0e^{-\frac{m\omega}{2\hbar}x^2}$, is the groundstate wave-function and the general solution is given by $\Psi_n(x) = A_n(a_+)^ne^{-\frac{m\omega}{2\hbar}x^2}$. Now, to determine the energy, we substitute the groundstate wavefunction to

$(a_+a_- + \frac{1}{2}\hbar\omega)\Psi(x) = E\Psi(x)$. Solving for E gives:

$$\begin{aligned}(a_+a_- + \frac{1}{2}\hbar\omega)\Psi_0(x) &= E\Psi_0(x)\\ \implies a_+a_-\Psi_0(x) + \frac{1}{2}\hbar\omega\Psi_0(x) &= E\Psi_0(x)\\ \implies a_+(0) + \frac{1}{2}\hbar\omega\Psi_0(x) &= E\Psi_0(x)\\ \implies E &= \frac{\hbar\omega}{2}\end{aligned}$$

Thus, giving the groundstate energy $E_0 = \frac{\hbar\omega}{2}$. The general solution for E is $E_n = (n + \frac{1}{2})\hbar\omega$. With this, we now have a way to determine the accuracy of the numerical models.

III. Finite Differences

Until now, we have solved the Schrödinger equation without any context for the scale we are working with. The three necessary units of measure are energy, mass, and position, particularly, energy is measured in mega-electron-volts (MeV), the position is calculated in Fermi's, and mass is represented as atomic mass units (amu).

#	Units	#	SI Units
1	MeV	$1.60218 * 10^{-13}$	J
2.0135	amu	$3.343 * 10^{-27}$	kg
1	Fermi	$1 * 10^{-15}$	m

Fig. 5: Conversion of units used in this study to SI units

The system of interest is the deuteron, specifically the interaction between the proton and neutron, with that Eq.

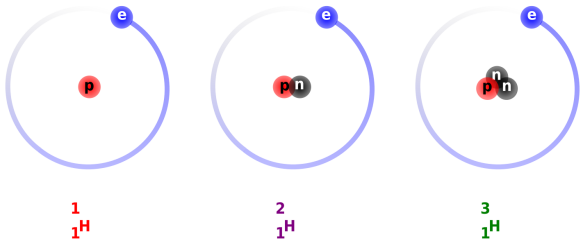


Fig. 6: The image to the left is the depiction of a typical hydrogen atom; followed by the deuteron, a hydrogen nucleus with an added neutron particle; and finally, the triton atom, which is a hydrogen nucleus with two added neutron particles being the heaviest hydrogen atom to exist.

(6), can be rearranged in the following manner:

$$\begin{aligned}E\Psi(x) &= -\frac{\hbar^2}{2m}\frac{d^2\Psi(x)}{dx^2} + V(x)\Psi(x)\\ \implies 0 &= -\frac{\hbar^2}{2m}\frac{d^2\Psi(x)}{dx^2} + (V(x) - E)\Psi(x)\\ \implies \frac{d^2\Psi(x)}{dx^2} &+ \frac{2m}{\hbar^2}(E - V(x))\Psi(x) = 0 \qquad (8)\end{aligned}$$

Since Eq. (8) is of the form $\frac{\partial^2 y}{\partial x^2} + K(x)y(x) = 0$ finite difference can be applied to find the energy of the system. Before performing the finite difference, some parameters are established as seen in Fig. (7) and (8). For this particular system, the potential to be solved is $V(x) = -60e^{-(x/1.65)^2}$. Now, suppose a, b are defined

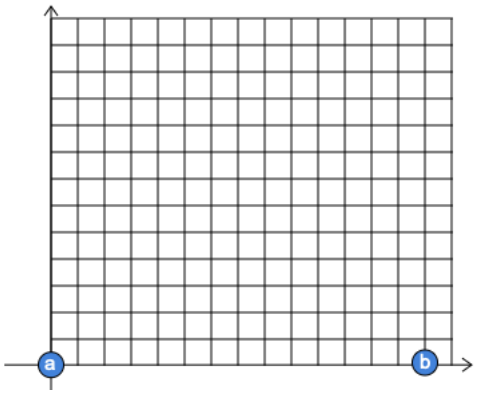


Fig. 7: An arbitrary domain in an interval $[a, b]$ with evenly discretized grids.

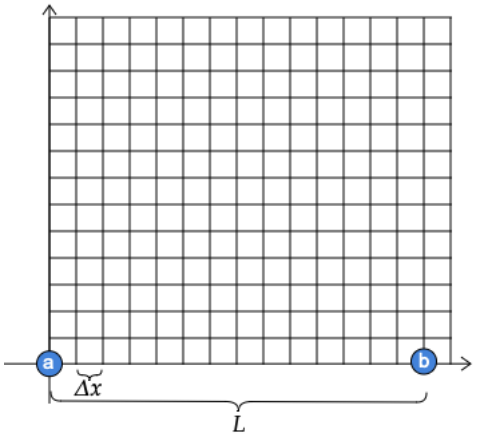


Fig. 8: With the respective discretization, $L := b - a$, $\Delta x := L/N$ where a chosen N determines the fineness of the grid.

as follows: $a := x_0$ where $a + \Delta x := x_1 \implies a + n * \Delta x = x_n$. Using this definition $\Psi_n := \Psi(x_n)$. If

https://digitalcommons.odu.edu/ourj/vol11/iss1/3
DOI: https://doi.org/10.25778/9dd7-vn32

$\Psi^{(n)}(x)$ is the nth derivative, a Taylor series expansion is expanded around Ψ_{n+1}, Ψ_{n-1} as follows:

$$\begin{aligned}\Psi_{n+1} &= \Psi(x) + \Delta x\Psi^{(1)}(x) + \frac{\Delta x^2}{2}\Psi^{(2)}(x) \\ &\quad + \frac{\Delta x^3}{6}\Psi^{(3)}(x) + \frac{\Delta x^4}{24}\Psi^{(4)}(x) + ... \\ \Psi_{n-1} &= \Psi(x) - \Delta x\Psi^{(1)}(x) + \frac{\Delta x^2}{2}\Psi^{(2)}(x) \\ &\quad - \frac{\Delta x^3}{6}\Psi^{(3)}(x) + \frac{\Delta x^4}{24}\Psi^{(4)}(x) + ... \\ \Psi_{n+1} + \Psi_{n-1} &= 2\Psi(x) + \Delta x^2\Psi^{(2)}(x) \\ &\quad + \frac{\Delta x^4}{12}\Psi^{(4)}(x) + O(\Delta x^6)\end{aligned}$$

Rearranging the last equation $\Psi_{n+1} + \Psi_{n-1}$ produces the scheme for the second-order derivative of the Schrödinger equation.

$$\Psi^{(2)}(x) = \frac{\Psi_{n-1} - 2\Psi(x) + \Psi_{n+1}}{\Delta x^2} - \frac{\Delta x^2}{12}\Psi^{(4)}(x) + O(\Delta x^4)$$

Essentially, what we get is the 4th-order Runge-Kutta (RK4) framework to solve the Schrödinger Equation. And, due to the nature of the equation, we can achieve a higher order accuracy by having a scheme that involves the 4th derivative.

Acting on the Schrödinger Equation with $1 + \frac{\Delta x^2}{12}\frac{d^2}{dx^2}$ gives:

$$\begin{aligned}(1 + \frac{\Delta x^2}{12}\frac{d^2}{dx^2})(\Psi^{(2)} + K(x)\Psi) &= \Psi^{(2)} + K(x)\Psi \\ &+ \frac{\Delta x^2}{12}\Psi^{(4)} + \frac{\Delta x^2}{12}\frac{d^2}{dx^2}(K(x)\Psi) = 0\end{aligned}$$

Substituting $\Psi^{(2)} + \frac{\Delta x^2}{12}\Psi^{(4)}$ back into the RK4 scheme from before produces:

$$\begin{aligned}0 = \frac{\Psi_{n-1} - 2\Psi(x) + \Psi_{n+1}}{\Delta x^2} + K(x)\Psi(x) \\ + \frac{\Delta x^2}{12}\frac{d^2}{dx^2}(K(x)\Psi(x)) + O(\Delta x^4)\end{aligned}$$

After some rearranging, the equation becomes:

$$\begin{aligned}\Psi_{n-1} - 2\Psi(x) + \Psi_{n+1} = \Delta x^2 K(x)\Psi(x) \\ + \frac{\Delta x^4}{12}\frac{d^2}{dx^2}(K(x)\Psi(x)) + O(\Delta x^6)\end{aligned}$$

Previously, it was shown for a function $f^{(2)}(x)$ the finite difference scheme is given by $\frac{f_{n-1} - 2f_n + f_{n+1}}{\Delta x^2}$. Substitute this for the $K(x)\Psi(x)$ term, and the result for the 6th order Numerov method follows.

$$\begin{aligned}\Psi_{n+1} = \frac{(2 - \frac{5}{6}\Delta x^2 K(x))\Psi(x) - (1 + \frac{1}{12}\Delta x^2 K_{n-1})\Psi_{n-1}}{1 + \frac{1}{12}\Delta x^2 K_{n+1}} \\ + O(\Delta x^6)\end{aligned}$$

We proceed with the method from both the right and left boundary, which will meet at an arbitrary point M.

$$\begin{aligned}\Psi^+ &= \Psi_{n+1} \\ &= \frac{(2 - \frac{5}{6}\Delta x^2 K_n)\Psi_n - (1 + \frac{1}{12}\Delta x^2 K_{n-1})\Psi_{n-1}}{1 + \frac{1}{12}\Delta x^2 k_{n+1}} \\ &\quad + O(\Delta x^6) \\ \Psi^- &= \Psi_{n-1} \\ &= \frac{(2 - \frac{5}{6}\Delta x^2 K_n)\Psi_n - (1 + \frac{1}{12}\Delta x^2 K_{n+1})\Psi_{n+1}}{1 + \frac{1}{12}\Delta x^2 k_{n-1}} \\ &\quad + O(\Delta x^6)\end{aligned}$$

With this specific potential applied to the equation, the conditions considered are:

$$\begin{gathered}\Psi(0) = 0, \quad \Psi(\Delta x) = \Delta x, \\ \Psi(b = 20) = e^{-k*(20)}, \ \Psi(20 - \Delta x) = e^{-k*(20 - \Delta x)}\end{gathered}$$

Where k is fixed by the Schrödinger Equation, $k^2 = 2mE/\hbar^2$, and any two arbitrary energies can be chosen for the model, which acts as a sort of guess, and a specific value was chosen to aid the convergence rate of the model.

$$E(0) = -5, \ E(1) = -5.005$$

We note that a specific meeting point M was also chosen for convergence reasons. The parameter indicates the iteration in which we performed the Numerov method. For $N = 1000$, after ten iterations, our solution converges to an energy $E = -2.29$ MeV with a tolerance of $4.889\text{E}{-13}(|E_{n\text{ step}} - E_{n-1\text{ step}}|)$, but the wave function looks wrong.

Before we compute the integral, we need to take care of the discontinuity at the meeting point M since Ψ is continuous. We know that

$$\begin{aligned}\hat{\Psi}(x_M) &\approx A\Psi_+ \\ \hat{\Psi}(x_M) &\approx B\Psi_-\end{aligned}$$

$\implies \Psi_+ = \frac{B}{A}\Psi_-$ at M. Where $\hat{\Psi}$ is denoted as follows:

$$\hat{\Psi} = \begin{cases}\Psi_+(x_i), & 0 < i \leq M \\ \Psi_-(x_i), & i > M\end{cases}$$

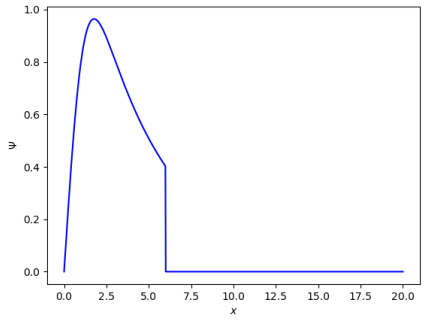


Fig. 9: Before matching at the meeting point $M = 300$.

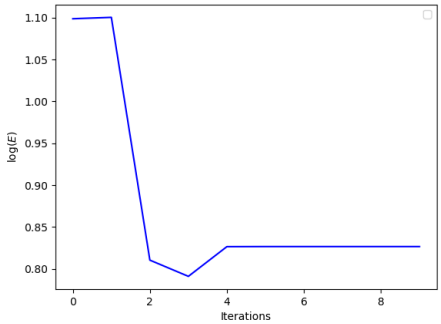


Fig. 10: The plot showing the convergence of the energy for each iteration.

We also use this quantity to find the energy for the next step, as the first two energies serve as a guess. Define Δ as follows:

$$\Delta := \left| \frac{\Psi'_+(M)}{\Psi_+(M)} - \frac{\Psi'_-(M)}{\Psi'_-(M)} \right|$$

where $\Psi'_\pm := \frac{\mp\Psi_\pm(M\mp1)\pm\Psi_\pm(M)}{\Delta x}$. Using this relationship, new energies can be calculated for the next steps as follows since we now have two guess energies:

$$\begin{aligned} m &:= \frac{\Delta_i - \Delta_{i-1}}{E_i - E_{i-1}} \\ \implies y - \Delta_i &= m(x - E_i) \\ y &= m(x - E_i) + \Delta_i \\ \implies E_{i+1} &= -\Delta_i/m + E_i \end{aligned}$$

After iterating the Numerov method, the algorithm converges to the energy previously discussed as seen in Fig (10). So we take points produced from $\hat{\Psi}$:

and use the Trapezoid method to integrate the new function. What we get is as follows:

$$\int_\Omega |\hat{\Psi}|^2 \approx \sum_{i=1}^{N} \left| \frac{\big(\hat{\Psi}(x_i) + \hat{\Psi}(x_i - 1)\big)\Delta x}{2} \right| = c^2$$

Finally, to get the wavefunction Ψ, which we desire, we divide $\hat{\Psi}$ by c producing the following result $\hat{\Psi}/c = \Psi$ fulfilling the normalization condition. Generally, when

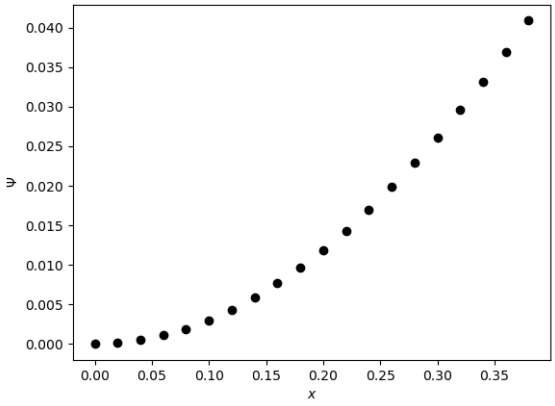


Fig. 11: The first 20 points of Ψ.

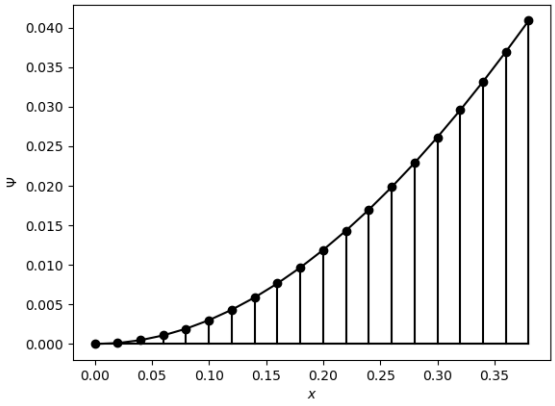


Fig. 12: The first 20 points of Ψ applying the Trapezoid method.

solving a quantum mechanical system, a normalizing proceedure must be fulfilled, namely

$$\int_\Omega |\Psi(x)|^2 = 1$$

When comparing to experimental data, we see that the ground state energy of the deuteron with the given potential has a $|\%\text{-error}| < 1$. Similarly, when using this method to solve the harmonic oscillator problem, we get

https://digitalcommons.odu.edu/ourj/vol11/iss1/3
DOI: https://doi.org/10.25778/9dd7-vn32

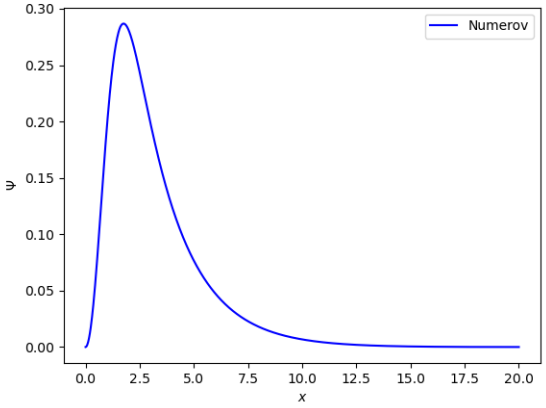


Fig. 13: Normalized Ψ for the deuteron.

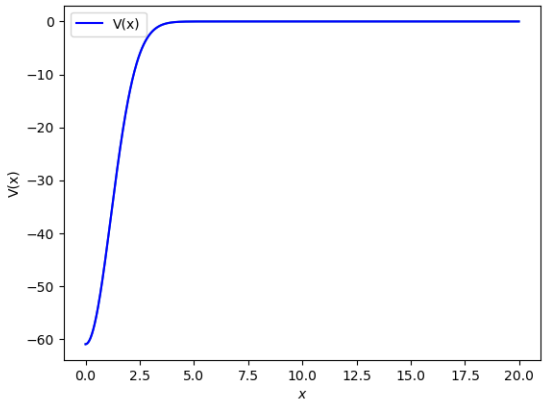


Fig. 14: Respective potential $V(x) = -60e^{-(x/1.65)^2}$.

a similar accuracy with an energy of 3.262 MeV using the parameters $\omega, m = 1$ compared to the theoretical value of 3.291 MeV.

IV. Neural Networks

We take advantage of the fact that Feed Forward Neural Networks(FNN)/Artificial Neural Networks(ANN) are universal approximators of any equation or system [9]. There are two steps for the neural network method. First, even before performing the forward propagation, the Metropolis- Hastings algorithm, which is a Markov Chain-Monte Carlo method [13], randomly samples from a probability distribution similar to Ψ and produces the trial wave function Ψ^* which as more sampling occurs, $\Psi^* \to \Psi$.

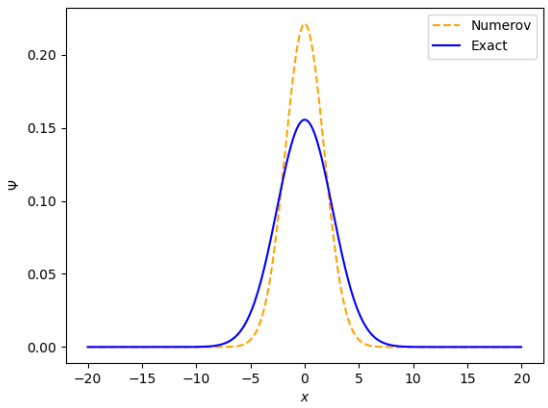


Fig. 15: The Ψ_0 produced by the Numerov Method compared to the exact solution for the Harmonic Oscillator. The reason for the overshoot is likely due to the method over compensating for the energy when we don't have the same result as the exact solution.

A. *Metropolis-Hastings*

The burnout period for the Markov Chain to approach the desired stationary state is about $50\% - 60\%$ of the total N samples taken, and the remaining samples are used in the Metropolis-Hastings algorithm. When the stationary state is reached, a desirable property arises where the Markov Chain is reversible. This means that the probability of a certain sample times the probability approaching a new sample is equal to the probability of the new sample times the probability from the new sample approaching the original sample. We did not prove the existence of stationary distributions, and only experimentally tested to see if the Markov Chain converged to some distribution.

$$\mathbb{P}(\theta_s) \cdot \mathbb{T}(\theta_s \to \theta_s') = \mathbb{P}(\theta_s') \cdot \mathbb{T}(\theta_s' \to \theta_s) \tag{9}$$

(10)

The Detailed Balance principle is the premise for the Reversible Markov Chain. In the above equation (9), $\mathbb{P}$ indicates the probability and $\mathbb{T}$ is a different notation used to represent the probability of transition from one sample to another sample.

For this study, the variables sampled are the various positions at which the deuteron is located. The other object also being sampled is the movement of the particles to a new position. These moving samples are "walkers," named this way because there is a certain probability for their movement closer to or further away from the

mean of the wave function. In this analysis, we assigned 400 walkers to various positions and sampled them until achieving a stationary state. Observe that the distance at which the particles move is based on a Gaussian distribution. These sampled data are not used until a stationary state is reached. After this burn-in period, we now use these artificially produced samples in the Metropolis-Hastings algorithm, to apply to the Monte-Carlo simulation. This approximates the probability using another Gaussian distribution that covers Ψ for faster convergence.

The Gaussian which the sample is subjected to is:

$$G(x) := \frac{2}{\sqrt{2\pi}} e^{-x^2/2}$$

To ensure that a somewhat true random sampling occurs,

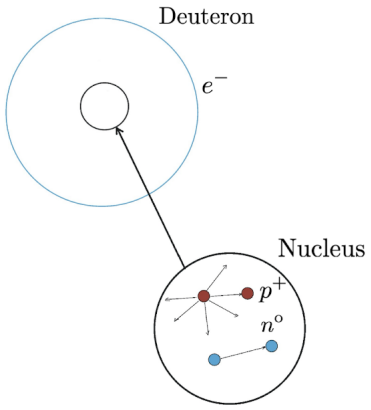


Fig. 16: This figure shows the representation of the modeled isotope. For the Metropolis-Hastings algorithm, there will be 400 different walkers moving around the nucleus, and the probabilities of these walkers are later sampled. A separate Gaussian distribution is independently used for the distance and direction to determine the new position. Consider the proton-neutron system, where it is calculated together as a single particle using a relative position.

only every ten samples of the relative position of the particles are used in the Metropolis-Hastings algorithm. Once a cycle of this algorithm is performed, the position is then inputted into the neural network.

Now, moving to the explaination of the neural networks, for clarity, an explanation of the loss function is done first. Recall the Schrödinger equation Eq. (4) is written as follows:

$$i\hbar \frac{\partial \Psi(r,t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 \Psi(r,t) + V(r,t)\Psi(r,t)$$

The Hamiltonian operator is defined like this:

$$\hat{H} := -\frac{\hbar^2}{2m}\nabla^2 + V(r,t)$$

where it acts on the object to the right of it. This means that the equation can be rewritten as follows, with the interpretation that the left hand side represents the energy eigenvalue of the equation.

$$\begin{aligned} E\Psi &= \hat{H}\Psi \\ \implies E &= \frac{\hat{H}\Psi}{\Psi} \\ \implies &= \frac{\Psi^*}{\Psi^*}\frac{\hat{H}\Psi}{\Psi} \\ &= \frac{\Psi^*\hat{H}\Psi}{|\Psi|^2} \end{aligned}$$

In Bra-Ket notation, it is written as:

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{11}$$

Since the Ψ is inherently a probability, the sum of all of its values will give the approximated value of the energy.

$$\begin{aligned} E &= \frac{\Psi^*\hat{H}\Psi}{|\Psi|^2} \\ \implies &\approx \int_\Omega \frac{|\Psi|^2}{\int_\Omega |\Psi|^2}\frac{\hat{H}\Psi}{\Psi} d\vec{r} \\ \implies &\approx \sum_{\vec{r}\in\mathbb{P}} \frac{\hat{H}\Psi}{\Psi} \end{aligned} \tag{12}$$

This quantity (11) will be used as the loss function in the neural network to solve for (9).

B. Feedforward Neural Network Architecture

We utilized a feedforward neural network with an input layer consisting of 3 nodes, two hidden layers, each comprising 16 nodes, and a single output node. The activation function used in the hidden layers is the hyperbolic tangent ($\tanh$) function, and the output layer uses a linear activation function suitable for regression tasks. We took inspiration from [11] and modeled the ANN to solve the deuteron.

1) Forward Pass: The mathematical formulation of the forward pass is as follows:

a) Input Layer to Hidden Layer 1: Let $\mathbf{x}$ be the input vector with three elements, $W^{(1)}$ be the weight matrix of size 16×3, and $\mathbf{b}^{(1)}$ be the bias vector of size 16 for the first hidden layer. The input to the first hidden layer is computed as:

$$\mathbf{z}^{(1)} = W^{(1)}\mathbf{x} + \mathbf{b}^{(1)}$$

https://digitalcommons.odu.edu/ourj/vol11/iss1/3
DOI: https://doi.org/10.25778/9dd7-vn32

Applying the $\tanh$ activation function gives:

$$\mathbf{a}^{(1)} = \tanh(\mathbf{z}^{(1)})$$

b) Hidden Layer 1 to Hidden Layer 2: Let $W^{(2)}$ be the weight matrix of size 16×16 and $\mathbf{b}^{(2)}$ be the bias vector of size 16 for the second hidden layer. The input to the second hidden layer is:

$$\mathbf{z}^{(2)} = W^{(2)}\mathbf{a}^{(1)} + \mathbf{b}^{(2)}$$

Applying the $\tanh$ activation function gives:

$$\mathbf{a}^{(2)} = \tanh(\mathbf{z}^{(2)})$$

c) Hidden Layer 2 to Output Layer: Let $W^{(3)}$ be the weight matrix of size 1×16 and $b^{(3)}$ be the bias for the output layer. The input to the output layer is:

$$z^{(3)} = W^{(3)}\mathbf{a}^{(2)} + b^{(3)}$$

The output, assuming a linear activation function for regression, is:

$$\hat{y} = z^{(3)}$$

2) Loss Function: The loss function used is the minimization of Eq. (11) the energy/expectation value itself, defined as:

$$\mathcal{L} = E = \min_{W} \frac{\langle \Psi_V(W)|\hat{H}|\Psi_V(W)\rangle}{\langle \Psi_V(W)|\Psi_V(W)\rangle}$$

where W are the weights of the neural networks and Ψ_V is the variational wave function.

C. Backward Pass

To perform backward propagation, we compute the gradients of the loss function with respect to the weights and biases. Essentially, what happens is we optimize the gradients with respects to the weights $W^{(n)}, n = \{1, 2, 3\}$ to minimize the energy i.e.

$$\text{Find } \frac{\partial \mathcal{L}}{\partial W^{(n)}} \ni: \min_{W} E$$

Bear in mind, that we cannot compute every single derivative as that would be computationally expensive, so in this analysis, we use the ADAM Optimizer to compute and optimize the gradients stochastically. The weights in the hidden layer are optimized to minimize the energy every time the neural network is backpropagated.

D. ADAM Optimizer

The ADAM optimizer was used to update the network parameters. ADAM combines the benefits of RMSProp and Momentum [10] by maintaining an exponentially decaying average of past gradients (m) and past squared gradients (v).

For each parameter θ:

a) Initialization:

$$m_t = 0, \quad v_t = 0$$

$$\beta_1 = 0.9, \quad \beta_2 = 0.999, \quad \epsilon = 10^{-8}, \quad \alpha = \text{learning rate}$$

b) Update at time step t: Compute gradients:

$$g_t = \frac{\partial \mathcal{L}}{\partial \theta}$$

Update biased first-moment estimate:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$$

Update biased second-moment estimate:

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$$

Correct bias in the first moment:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$$

Correct bias in the second moment:

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$

Update parameters:

$$\theta_t = \theta_{t-1} - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

E. Results

We applied the neural network onto the deuteron with the same potential Fig. (14) as described in the previous section. After running the neural networks for 400 iterations with a learning rate of $\alpha = 1 * 10^{-4}$ we can see the energy decreases and is minimized by the optimizer, as seen in Fig. (17). It converges to

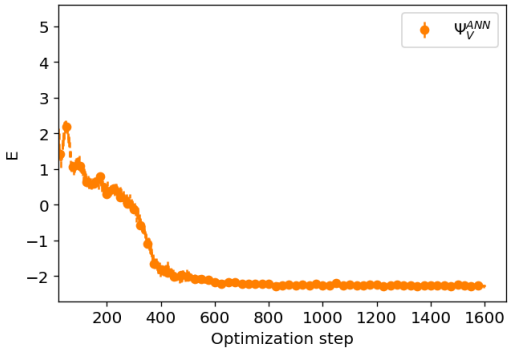


Fig. 17: The energy computed after each iteration.

around -2.266 MeV with a percent error that is $|\%\text{-error}| = 1.04$. To produce the wave function, since the model stochastically samples the position of the deuteron, we use the sampled data and produced a distribution of the measured positions of the particles. This distribution is the wave function as in Fig. (19).

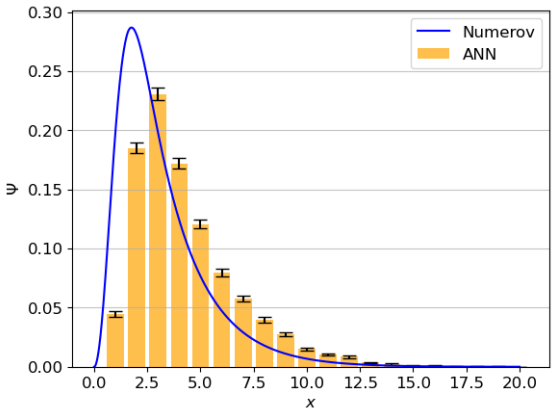


Fig. 18: A histogram showing the position sampled as a distribution of the deuteron, compared to the wave function computed from the finite difference.

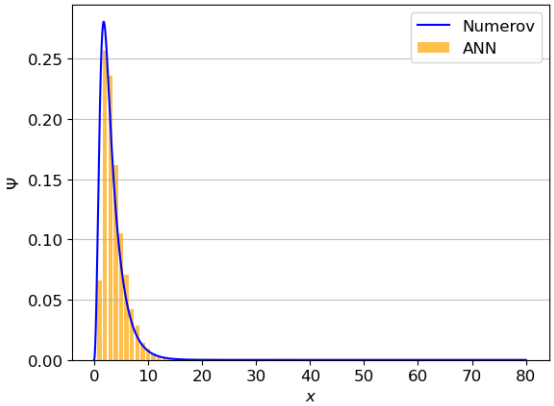


Fig. 19: A histogram showing the position sampled as a distribution of deuteron, compared to the wave function computed from the finite difference for 40 Fermi's showing increased accuracy of the model.

V. Conclusion

In this article, we compared three ways to analyze the Schrödinger equation and applied different potentials. The trend for each model is that it shows an increased capability compared to the previous model. The analytical method is limited to certain problems compared to the Numerov method with an increased ability to solve a larger array of systems by taking advantage of the computing power of modern computers. Ultimately, we see neural networks can solve even more difficult problems by utilizing artificial intelligence algorithms to solve problems in which finite difference is weak, such as many-body problems.

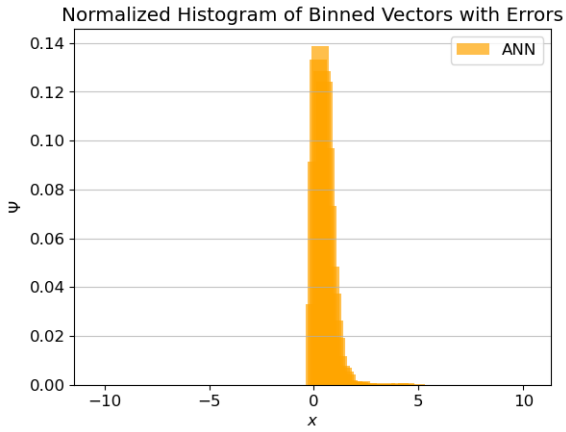


Fig. 20: A histogram of the harmonic oscillator using the neural networks.

In our analysis, we solved a two-body problem with the neural network and still achieved a relatively accurate energy compared to the exact and experimental solutions. The wave functions maintained a shape similar to that of the actual Ψ as well. What is also noteworthy is that the neural network was able to solve the Schrödinger equation with only the information provided from the Hamiltonian. This shows how much information is retained in this operator and that by taking advantage of it, an even more accurate and stable model can solve even more difficult problems.

We want to note that the ADAM optimizer used in this paper is a more general algorithm that can be swapped for the stochastic reconfiguration method to handle the complexities of the Schrödinger equation [15]. We did not account for all three dimensions, spin, and asymmetry of the particles to name a few of the properties that the neural network and stochastic reconfiguration are capable of. Because of the design of the problem in this analysis, the symmetry of the particles was accounted and it was not necessary to consider. The deuteron was solved radially making the system symmetrical to begin with.

Future directions for this analysis are implementing more appropriate optimizers for the problem such as the aforementioned stochastic reconfiguration method [11], [14]. An even more efficient method is the so-called Hessian accelerated stochastic reconfiguration method [2]. Extensions of this investigation can also include analyzing the convergence, stability, and also higher energy states of our methods and models. These

https://digitalcommons.odu.edu/ourj/vol11/iss1/3
DOI: https://doi.org/10.25778/9dd7-vn32

methods can also be applied to solve a wider array of problems that are at the intersection of interdisciplinary studies such as physics, biology, chemistry, etc. [1], [8], [16], [18]. To achieve an even more accurate and faster convergence of our solution, the implementation of physics-informed neural networks can also be used. This idea builds physics facts into the neural network model to assist the propagation of the weights and allow for the optimizers to minimize the loss functions more efficiently.

REFERENCES

[1] Vi D. Ao, Duy V. Tran, Kien T. Pham, Duc M. Nguyen, Huy D. Tran, Tuan K. Do, Van H. Do, and Trung V. Phan. A Schrödinger Equation for Evolutionary Dynamics. *Quantum Reports*, 5(4):659–682, December 2023. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[2] Claudio Attaccalite. RVB phase of hydrogen at high pressure: towards the first ab-initio Molecular Dynamics by Quantum Monte Carlo. October 2005. Publisher: SISSA.

[3] Robert Dicke and J Wittke. *Introduction to Quantum Mechanics*. Addison-Wesley, 1966.

[4] Johan Falk. *Students' depictions of quantum mechanics: a contemporary review and some implications for research and teaching*. PhD thesis, Uppsala University, Sweden, 2007.

[5] Jonathan Gorard. Deriving the time-independent Schrödinger equation. *Physics Education*, 51(6):063003, September 2016. Publisher: IOP Publishing.

[6] David J. Griffiths and Darrell F. Schroeter. *Introduction to Quantum Mechanics*. Cambridge University Press, 3rd edition, August 2018. ISBN: 9781316995433 Publisher: Cambridge University Press.

[7] David Hanson, Erica Harvey, Robert Sweeney, Theresa Zielinski, Mark Tuckerman, and Seymour Blinder. The Schrödinger Equation.

[8] Heiko Hergert. A Guided Tour of ab initio Nuclear Many-Body Theory. *Frontiers in Physics*, 8, October 2020. Publisher: Frontiers.

[9] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, January 1989.

[10] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization, January 2017. arXiv:1412.6980 [cs].

[11] Alessandro Lovato, Corey Adams, Giuseppe Carleo, and Noemi Rocco. Hidden-nucleons neural-network quantum states for the nuclear many-body problem. *Physical Review Research*, 4(4):043178, December 2022. Publisher: American Physical Society.

[12] T. Mart. How do I introduce Schrödinger equation during the quantum mechanics course? *Physics Education*, 56(2):025012, March 2021. arXiv:2010.15589 [hep-ph, physics:nucl-th, physics:physics, physics:quant-ph].

[13] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, June 1953.

[14] Chae-Yeun Park and Michael J. Kastoryano. Geometry of learning neural quantum states. *Physical Review Research*, 2(2):023232, May 2020. arXiv:1910.11163 [cond-mat, physics:quant-ph, stat].

[15] Sandro Sorella and Luca Capriotti. Green function Monte Carlo with stochastic reconfiguration: An effective remedy for the sign problem. *Physical Review B*, 61(4):2599–2612, January 2000. Publisher: American Physical Society.

[16] Huu S. Tieu and Martin F. Loeffler. Life Is Quantum Biology Effects Explained from the Schrödinger Equation, Serious or Life-Threatening Conditions or Diseases, and COVID-19 Results. *Journal of Biosciences and Medicines*, 10(10):164–172, October 2022. Number: 10 Publisher: Scientific Research Publishing.

[17] Steven Weinberg. *Lectures on Quantum Mechanics*. Cambridge University Press, Cambridge, 2 edition, 2015.

[18] Jun Yang. Making quantum chemistry compressive and expressive: Toward practical ab-initio simulation. *WIREs Computational Molecular Science*, 14(2):e1706, 2024. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1706.

APPENDIX

A. Momentum Optimizer

The Momentum optimizer is designed to accelerate the gradient descent algorithm by accumulating a velocity vector in directions of persistent reduction in the loss function. This helps to dampen oscillations and speed up convergence.

a) Mathematical Formulation: Given a parameter θ, the gradient of the loss function $\mathcal{L}$ with respect to θ at time step t is denoted as $g_t = \nabla_\theta \mathcal{L}(\theta_t)$.

The Momentum update rule is defined as:

$$v_t = \beta v_{t-1} + (1-\beta) g_t$$

$$\theta_t = \theta_{t-1} - \alpha v_t$$

where:

- v_t is the velocity at time step t.
- β is the momentum coefficient (typically $\beta = 0.9$).
- α is the learning rate.

The term v_t acts as an aggregate of past gradients, allowing the optimizer to maintain the direction of previous gradients while smoothing out noise.

B. RMSProp Optimizer

RMSProp (Root Mean Square Propagation) is an adaptive learning rate method that aims to resolve the issue of diminishing learning rates in the Adagrad algorithm by maintaining a moving average of the squared gradients.

a) Mathematical Formulation: For a parameter θ, the gradient at time step t is $g_t = \nabla_\theta \mathcal{L}(\theta_t)$.

The RMSProp update rule is:

$$E[g^2]_t = \beta E[g^2]_{t-1} + (1-\beta) g_t^2$$

$$\theta_t = \theta_{t-1} - \frac{\alpha}{\sqrt{E[g^2]_t + \epsilon}} g_t$$

where:

- $E[g^2]_t$ is the exponentially decaying average of past squared gradients.
- β is the decay rate (typically $\beta = 0.9$).
- α is the learning rate.
- ϵ is a small constant to prevent division by zero (typically $\epsilon = 10^{-8}$).

RMSProp adjusts the learning rate for each parameter individually based on the historical magnitudes of gradients, which allows for larger updates for infrequent and smaller updates for frequent parameters.

C. Combining Momentum and RMSProp: ADAM Optimizer

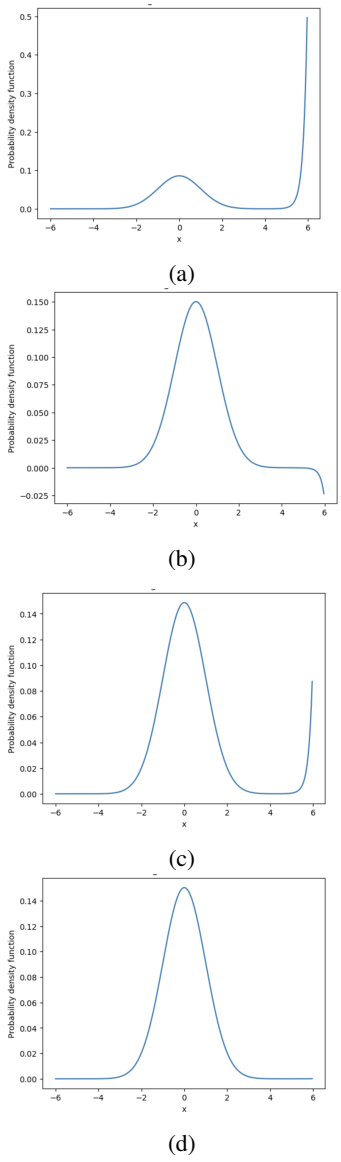
The ADAM (Adaptive Moment Estimation) optimizer combines the advantages of both Momentum and RMSProp. It maintains an exponentially decaying average of past gradients (m_t) and past squared gradients (v_t), and computes bias-corrected estimates to adapt the learning rate for each parameter.

The update rules for ADAM are provided in the main methods section of this document.

D. Note on the Numerov Method and usage of units

There is a different variation of the Numerov Method that instead of computing the energy as described above, the guess energy is iteratively reduced to a point where the boundary conditions are met. There is no matching in this variation instead, the left boundary is fixed to 0 for the harmonic oscillator potential and the right boundary is measured till $\lim_{x\to\infty} \Psi(x) = 0$. The following figures shows this phenomena.

For simplicity and proof of concept, the the harmonic oscillator used 1 as a simple parameter to achieve the result presented here for both the mass and frequency. The MeV$\cdot s$ reduced Planck constant was used for the harmonic oscillator and the units used for the deuteron were MeV for mass, MeV·Fermi for the reduced Planck constant, Fermi for distance, and MeV for energy. Actual quantities of the hydrogen isotope was used for the Numerov method and Neural Networks.



(a)

(b)

(c)

(d)

Fig. 21: As the energy gets closer and closer to the true value, the right boundary gets closer to 0.

https://digitalcommons.odu.edu/ourj/vol11/iss1/3
DOI: https://doi.org/10.25778/9dd7-vn32