Stochastic Mechanics and Neural Networks

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

der Naturwissenschaftlichen Fakultät II
Chemie, Physik und Mathematik

der Martin-Luther-Universität
Halle-Wittenberg

vorgelegt von

Herrn Kai-Hendrik Henk

Gutachter: Prof. Dr. Wolfgang Paul
Prof. Dr. Steffen Trimper
Prof. Dr. Bernd Rosenow

Datum der öffentlichen Verteidigung: 25.11.2025

”Well I suppose now is the time for me to say something profund.

...

Nothing comes to mind, let’s do it.”

Jack O’Neill - Stargate-SG1

Zusammenfassung

Die Quantenmechanik ist eine der erfolgreichsten Theorien der modernen Physik. Ihre Interpretation ist jedoch immer noch umstritten und viele Interpretationen werden heute öffentlich diskutiert. Die Nelsonsche stochastische Mechanik bietet einen alternativen Ansatz zur nichtrelativistischen Quantenmechanik auf der Grundlage der Newtonschen Mechanik. Die Quanten-Hamilton-Gleichungen, abgeleitet aus dem Quanten-Hamilton- oder Pavon-Prinzip unter Verwendung der stochastischen Theorie der optimalen Steuerung, können zur Berechnung der osmotischen Geschwindigkeitsfelder mit der Bender/Steiner-Methode verwendet werden. Dieses Geschwindigkeitsfeld enthält alle Informationen dieser Beschreibung. Die Bender/Steiner-Methode jedoch rechnerisch ineffizient, schwer verständlich und schwer zu programmieren ist.

In dieser Arbeit leiten wir das stochastisch mechanische Äquivalent zum Rayleigh-Ritz-Prinzip der Quantentheorie her. Dieses Prinzip wird anschließend verwendet, um einen genetischen Algorithmus zu erstellen, der den quantenmechanischen Grundzustand berechnet, also die osmotische Geschwindigkeit zusammen mit der Grundzustandsenergie berechnet, indem zunächst das abgeleitete Energiefunktional minimiert und dann eine Riccati-Gleichung gelöst wird. Die osmotischen Geschwindigkeiten werden mithilfe neuronaler Netze parametrisiert. Der Algorithmus wird an drei eindimensionalen Potentialen getestet, wobei die Grundzustände und angeregten Zustände durch Berechnung der Grundzustände der supersymmetrischen Partner-Hamilton-Funktionen berechnet werden.

Der in dieser Arbeit entwickelte Algorithmus wird dann verwendet, um den Grundzustand zweier Tweezerpotentiale zu lösen. Diese werden unter Verwendung von Methoden der Zeitreihenanalyse miteinander und mit dem harmonischen Oszillator verglichen. Zusätzlich verwenden wir die Itô-Formel, um zwei gekoppelte stochastische Differentialgleichungen herzuleiten, was eine Phasenraumbeschreibung der Quantenmechanik liefert, ohne die Heisenbergsche Unschärferelation zu verletzen. Diese Beschreibung wird verwendet, um die Heisenberg-Unsicherheit des Grundzustands der beiden Tweezerpotentiale und des harmonischen Oszillators, ihre Phasenraumporträts sowie Phasenporträts der kohärenten Zustände des harmonischen Oszillators zu berechnen.

Abstract

Quantum mechanics is one of the most successful theories of modern physics. Its interpretation however is still controversial with many interpretations being discussed in public today. Nelson's stochastic mechanics description offers an alternative approach to non-relativistic quantum mechanics based on Newtonian mechanics. The quantum Hamilton equations, derived from the quantum Hamilton or Pavon principle using stochastic optimal control theory can be used to calculate the osmotic velocity fields using the Bender/Steiner method. That velocity field contains all the information in this description. The Bender/Steiner method is however computational inefficient, difficult to understand, and to program.

In this thesis, we derive the stochastic mechanical equivalent to the quantum-mechanical Rayleigh-Ritz principle. This principle is then used to build a genetic algorithm that determines the quantum mechanical ground state, ergo calculates the osmotic velocity together with the ground state energy by first minimizing the derived energy functional and then solving a Riccati equation. The osmotic velocities are parametrized by using neural networks. The algorithm is tested on three one-dimensional potentials, calculating their ground states and their exited states by computing the ground states of the supersymmetric partner Hamiltonians.

This new and efficient algorithm is then used to solve the ground state of two tweezer potentials used in levitodynamics. These are then compared to each other and with the harmonic oscillator by using methods from time series analysis. Additionally, we use the Itô-formula to derive two coupled stochastic differential equations, which yields a phase space description of quantum mechanics without violating the Heisenberg uncertainty principle. This description is applied to calculate the Heisenberg uncertainty of the ground state of the two tweezer potentials and the harmonic oscillator, their phase space portraits as well as phase portraits of the coherent states of the harmonic oscillator.

Contents

Chapter 1

Introduction

The advent of quantum mechanics at the beginning of the 20th century lead to an explosion of research concerning microscopic systems. The theoretical works of Planck [Pla01], Einstein [Ein05], Schrödinger [Sch26a; Sch26b; Sch26c; Sch26d; Sch26e], Heisenberg [Hei27], Max Born and Pascal Jordan [BJ25; Bor26], John von Neumann [Von32], Nils Bohr [Boh05; Boh13a; Boh13c; Boh13b], Paul Dirac [Dir81; DF26] together with experimental work of Louis Alvarez [AB40; Hab07], Ernest Rutherford [Rut12], Otto Stern and Walther Gerlach [GS22], Robert Milikan [Mil10], J.J. Thomson [Tho95; Tho98], Davisson and Germeer [DG27] lead to a new understanding of the universe. Contrary to classical physics, what physics before 1900 came to be called, one could no longer predict and measure trajectories with arbitrary accuracy. On small scales, one could only predict probabilities of measurements. Quantization explained a lot of previously unexplainable phenomena like the spectra and stability of atoms as well as the frequency distribution of black body radiation.

Quantum mechanics also lead to new developments in technology. The theory of semiconductors lead to the rise of electronics and with it computers, the internet and smartphones. Lasers have become a ubiquitous technology, having given rise to new research methods, uses in medicine and engineering among other things. Understanding of the inner workings of nuclei brought about the adoption of nuclear magnetic resonance as a diagnostic method in medicine and as a research method. This thorough understanding of the nucleus also lead to the utilization of nuclear fission in nuclear power plants and nuclear weapons. Nuclear fusion, touted by some as the energy source of the future, has only been applied in thermonuclear weapons up until now, with uses in fusion reactors not yet getting past the experimental research phase.

Even though quantum mechanics is one of the most successful theories of physics, being thoroughly tested and applied in many ways, the interpretation of quantum mechanics was controversially discussed. Many questions were raised by quantum mechanics, like ”Where does classical physics end and quantum mechanics start?” and ”What is the wave function and is it real (in a metaphysical context)?” had to be answered. The answers to these questions distinguish the many interpretations of quantum mechanics.

The most popular, due to it being *the* way quantum mechanics is being taught in universities around the globe, is the Copenhagen interpretation, named after the place of work of its biggest supporter, Nils Bohr. The particle obeys the Schrödinger equation, ergo is described by the wave function, until it interacts with a classical object. Then, the wave function collapses, something that is not predicted by the Schrödinger equation and becomes

a classical object and is measured. Many measurements yield a distribution of measurements as predicted by the Schrödinger equation. This interpretation has a lot of issues. One is its anti-realism. The wave function is not something real but only contains all knowable information of the quantum systems. Whatever happens before the measurement is not knowable. Quantum mechanics thus does not describe nature but only predicts measurement outcomes. So, the measurement becomes the central aspect of the Copenhagen interpretation. This point of critique was already made by Schrödinger with his famous thought experiment of Schrödingers cat. In the Copenhagen interpretation, a cat in a closed box with a death trap that is triggered by a quantum mechanical process is simultaneously dead and alive. This absurdity is commonly used to explain quantum mechanics to laypersons. A proponent would defend this absurdity by claiming the classical nature of the cat. This however only points to the next problem of the Copenhagen interpretation. The classical and quantum realms are strictly but not properly divided, which is also called the Heisenberg cut. What makes a detector classical and a particle quantum until it reacts with the detector? And is the detector not made of atoms that have to obey quantum mechanics? These questions cannot be answered by the Copenhagen interpretation.
Eugene Wigner criticized the role of the measurement process in the Copenhagen interpretation with his thought experiment of Wigner's friend [Wig95]. Consider a measurement of a quantum system. The wave function collapses for the researcher and he makes an observation. His friend however is outside the room. While the researcher makes an observation, for his friend the researcher and the quantum system are entangled, ergo the wave function of the quantum system did not yet collapse for the friend of the researcher. Only when the friend enters the room and asks the researcher, what observation he measured, the wave function would collapse from the friends point of view. Thus, a paradox occurs with two different times at which the quantum system's wave function collapses. All of those contradictions and paradoxes lead to many less commonly known interpretations.
To explain the ad-hoc loss of quantum-ness at measurement, many objective-collapse theories were proposed over the years. They search to formulate a realistic theory of wave function collapse. The most famous of those is the Diósi–Penrose model, proposed by Lajos Diósi [Dio87] and Roger Penrose [Pen96] which predicts gravitationally induced wave function collapse. This theory makes predictions on collapse times which have not been experientially verified. Related to objective-collapse theories is the von Neumann–Wigner interpretation [WM67; LBL39; PE97], that proposes that consciousness is necessary to make the wave function collapse, to conclude a measurement. It argues, that a detector is part of the physical world and thus has to obey quantum mechanics. The human mind however is the external observer and thus cannot be treated quantum mechanical. Thus, consciousness collapses the wave function. This theory however only delays the measurement from the detector to the researcher. Is consciousness a measurable quantity, and if yes, how much is deeded to collapse a wave function? And what even is consciousness? Thus, the only thing this theory accomplishes is a replacement of physical questions concerning the measurement process with philosophical questions regarding the nature of consciousness and the human mind.
Another work around of the wave function collapse is the many-worlds theory. First proposed by Hugh Everett III [Eve57], it explains the wave function collapse by not having one at all. If a quantum particle interacts with a detector, entanglement occurs. The detector entangles with the particle and then with the researcher operating the detector and so on and so on. The researcher only sees one of the possible outcomes of the measurement, because he experiences

only one part of the wave function that now describes him, the detector and the particle. In another part of the wave function, another world in terms of Everett, he experiences another possible measurement. Of course this entanglement does not stop with the researcher and did not start with that particle. The entire universe is a shared wave function of all particles describing ever larger branching upon each interaction of two particles. This means that the predictions of this theory are congruent with the ones predicted by the Schrödinger equation without the need for the ad-hoc, unmathematical collapse of the wave function. The many worlds however cannot communicate with each another, thus it is not possible to put this theory to test. (There exists a thought experiment called quantum suicide or quantum immortality [Teg07], in which a researcher would place himself into an experiment similar to the Schrödinger's cat experiment only to lower the probability of survival to a minimum. If the researcher would however exit this contraption alive, many-worlds proponents believe, this would be prove their favored interpretation.)
There are of course many more interpretations like, without any claim to completeness, quantum Bayesianism, also called QBism [Hea23], relational quantum mechanics [Rov25] or modal quantum mechanics [LD24]. The last class of interpretations that will be discussed here are the hidden-variable theories. The most prominent of the hidden-variable theories is the de Broglie-Bohm or pilot-wave theory. Here, particles exist even if unobserved and have trajectories that are guided by a pilot wave. This pilot wave obeys the Schrödinger equation. The particle trajectory is given by a differential equation which depends on the pilot wave. Thus, the wave function does not have to collapse for a measurement to occur. However, the measurements proposed by this theory are the same as predicted by the Schrödinger equation. Hidden variable theories are commonly disregarded due to a misunderstanding of the violation of Bell's inequalities.

In their 1935 paper [EPR35], Albert Einstein, Boris Podolsky and Nathan Rosen described a thought experiment of two entangled particles. If the two particles are then placed far away from each other and one is measured, it would instantaneously influence the other particle, violating the cosmic speed limit of the speed of light. This they called "spooky action at a distance" and they deduced, that quantum mechanics must be incomplete. Based on this thought experiment, David Bohm proposed an experiment of two entangled particles with spin $\frac{1}{2}$. If the correlation of the measurement as a function of the angle disobeys an inequality, now called Bell's theorem or Bell's inequality [Bel64], nature is non-local. A similar experiment with photons was later conducted and resulted in the experimentalists, Alain Aspect, John F. Clauser and Anton Zeilinger receiving the Nobel price of physics in 2022. There is however a common misconception, that the violation of Bell's inequality disproves all hidden-variables theories. This misunderstanding is even propagated by the Nobel committee [Rel22]. The violation of Bell's inequality only proves the impossibility of local hidden-variable theories [MGS24]. The pilot-wave theory is however non-local and its disregard for this reason premature.

Another hidden-variable theory and the focus of this thesis is Nelson's stochastic mechanics interpretation of quantum mechanics. First proposed by Edward Nelson in 1966 [Nel66] to derive the Schrödinger equation from Newtonian mechanics. Nelson proposed that microscopic particles undergo a conservative Brownian motion with Newtons second law being obeyed in the mean. With this he derived the Madelung equations [Mad26], which are a

hydrodynamic formulation of the Schrödinger equation, which are equivalent in the ground state of quantum systems. This was an attempt to give a realistic description of quantum mechanics [Nel85]. It has a big advantage in its relation to time. Time predictions can naturally be made in this interpretation, even though quantum mechanics has no time measurement operator. Thus Nelson's stochastic mechanics approach yield new predictions which are in agreement with experiments. However, the treatment of quantum mechanical systems with this interpretation relied on first solving the Schrödinger equation and from the resulting wave function deriving the velocity fields needed for the calculation of the stochastic trajectories.
Michele Pavon derived a variational principle similar to the Hamilton principle in classical mechanics [Pav95], thus called the quantum Hamilton principle or Pavon principle.
In 2017, Jeanette Köppe et al. derived from this principle the quantum Hamilton equations [Kö18; KGP16], finalizing a family of formalisms for quantum mechanics similar to the different formalism of classical mechanics. The quantum Hamilton equations offer an approach of calculating the velocity fields without relying on the Schrödinger equation. The method implemented was introduced by Christian Bender and Jessica Steiner [BS12]. The two quantum Hamilton equations, coupled stochastic differential equations, were iteratively integrated, one forward and one backward in time. The resulting velocity fields are the Nash equilibrium, solving the problem of having to rely on the Schrödinger equation.
Michael Beyer et al. derived in 2023 [BP23; Bey23] a way to describe particle spin in stochastic mechanics by extending it to curved manifolds and employing Euler angles to describe spin using the $\mathfrak{so}(3)$ rotation group. This resulted in quantum Hamilton equations for all three Euler angles with their conjugate momentum. The velocity fields could be obtained numerically with the Bender/Steiner algorithm reproducing the Stern-Gerlach experiment. But more importantly, using two simultaneous Stern-Gerlach experiments, it was shown, that the stochastic mechanics approach to quantum mechanics also violates Bell's inequality, once and for all disproving that Bell's theorem rules out all hidden-variable theories.
Coming back to the origin of the different interpretations, the question of the nature of the wave function is solved by not needing it in the first place. The particle was always a particle. The classical realm and the quantum realm are also not strictly divided. There is a continuous transition from the quantum domain to the classical domain with growing mass of the particles. The larger the mass of a particle, the less influence of the stochastic motion exists, the "more deterministic" the particle behaves.
Nelson's stochastic mechanics has also a few shortcomings. Aside from a few mathematical concerns regarding stochastic quantization, raised by Timothy C. Wallstrom in 1994 [Wal94], the biggest disadvantage of this interpretation is the practicability. The Bender/Steiner algorithm is slow, inefficient and suffers greatly from the curse of dimensionality. One could of course solve the Schrödinger equation. This is however conceptually very dissatisfying, having to rely on other interpretations of quantum mechanics to do anything practical in this interpretation.

Neural networks have been used in the sciences for a long time. Their almost universal applicability is due to them being almost universal function approximators, only being limited by computational restrictions due to limited resources.
After being proposed by Kolmogorov [Kol57] in 1957 as a solution to Hilbert's 13th of his 23 [Hil02] problems and the Dartmouth Conference of 1956, they gained ever more popularity in the fifties and sixties. After a short slowdown in applications, they resurfaced in the

1980s. In 1985, John Hopfield proposed the eponymous Hopfield networks [Hop82; Hop84] which he used together with David W. Tank to find a solution of the traveling salesman problem [HT85]. One year later, it was shown, that backpropagation is a valid method to use neural networks for practical regression problems [RHW86], giving the foundation for deep learning.
In 1990, neural networks were started to be used at CERN for analysis and simulation of particle collider experiments [Gle18]. In 1996, IBM's Deep Blue beat Garry Kasparov in chess [CHH02]. Then in 2012, convolutional neural networks were used [KSH12] to win the ImageNet 2012 challenge [Car25]. This resulted in a boom of applications of neural networks. DeepMind used neural networks in 2021 to solve the protein folding problem [JMJ21], resulting in a Nobel price in 2024 [Rel24a]. In the same year John Hopfield received a Nobel price [Rel24b] together with Demis Hassabis and John Jumper for their work on Hopfield networks and restricted Boltzmann machines, respectively. It thus makes sense to use neural networks for parametrization to solve physical problems.

There are many methods of finding the optimal weights and biases of neural netowrks. Genetic algorithms offer another approach to use neural networks. First proposed by Alan Turing in 1950 [TUR50], they were studied by many different scientists with the aim of finding a good optimization tool which took inspiration from evolution. In the 1960s, genetic algorithms as we know them now were invented by John Holland(e.g. [Hol62]). The original goal was to simulate natural occuring adaptation. The definitve framework of genetic algorithms was published by him in 1975 [Hol75].
Genetic algorithms were used for example in accelerator physics to find the optimal parameters of an accelerator beam [BPC16]. They are also a staple in bioinformatics, financial mathematics and engineering.

This thesis will first give an introduction to stochastic processes and their underlying mathematics. Then the stochastic mechanics approach to quantum mechanics will be discussed together with its connections to classical and conventional quantum mechanics. The fourth chapter will give a conceptual introduction of genetic algorithms. The fifth chapter is the first of two main parts of this thesis. First the stochastic mechanics equivalent of the Rayleigh-Ritz variational principle will be derived. Then, a genetic algorithm will be developed. As a proof of concept this algorithm will be tested on three different potentials, followed by a discussion of the parameters of this algorithm. The sixth chapter will further inquire the advantages of Nelson's interpretation of quantum mechanics. First, the use of time series analysis will be shown by treating a particle subject to three different potentials and comparing their properties to one another in order to gain valuable insight for experimental uses. Then a method is derived to gain sample paths in phase space that nonetheless obey the Heisenberg uncertainty principle. Finally, the thesis is concluded with a summary of all the results and with an overview of open questions and an outlook to future research.

Chapter 2

Stochastic Calculus

To get a thorough understanding of the description of quantum mechanics as conservative Brownian motion [Nel66], one must first understand stochastic processes. This chapter aims at giving a basic mathematical description of the fundamental concepts of stochastic analysis [Kö18; Gar09].

2.1 Random Variables and Stochastic Processes

A stochastic process $(X(t))_{t\in I}$ is a family of $\mathbb{R}^d$ (or complex-valued) random variables with an index set I [Kö18; Arn73]. To signal that X_t is a stochastic process, one commonly writes X_t instead of $X(t)$. Random variables are described by a sample space Ω, all possible events, $\mathcal{F}$ a subset of Ω which denotes the outcomes we want to study and a probability measure $\mathcal{P} : \mathcal{F} \to [0, 1]$ which gives the probabilities of certain events, which all obey a Borel-σ-Algebra, which means

- $\Omega \in \mathcal{F}$
- $\mathcal{F}$ is closed under complementation, ergo if $A \in \mathcal{F}$, then $A^c := \Omega \setminus A \in \mathcal{F}$
- $\mathcal{F}$ is closed under countable unions, ergo if $A_i \in \mathcal{F}$ then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$

Usually $\mathcal{P}$ is given by a probability distribution $\rho_X(A) = \mathcal{P}(A)$ such that for a random variable $X : \Omega \to \mathbb{R}^d$ the expectation is

$$\mathbb{E}[X] = \int_\Omega X(\omega)\mathrm{d}P(\omega) = \int_{\mathbb{R}^d} x\rho(x)\mathrm{d}x \tag{2.1}$$

with X being said to be integrable if $\mathbb{E}[|X|] < \infty$. The variance is defined as

$$\mathrm{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2. \tag{2.2}$$

If f is a function of x, then its mean is

$$\mathbb{E}[f(X)] = \int_\Omega f(X(\omega))\mathrm{d}P(\omega) = \int_{\mathbb{R}^d} f(x)\rho(x)\mathrm{d}x. \tag{2.3}$$

One of the most important stochastic processes is the Wiener process. In 1827, Robert Brown examined pollen in water under a microscope and observed irregular motion. Norbert Wiener described that motion with a stochastic process, the Wiener process W_t, with the following properties:

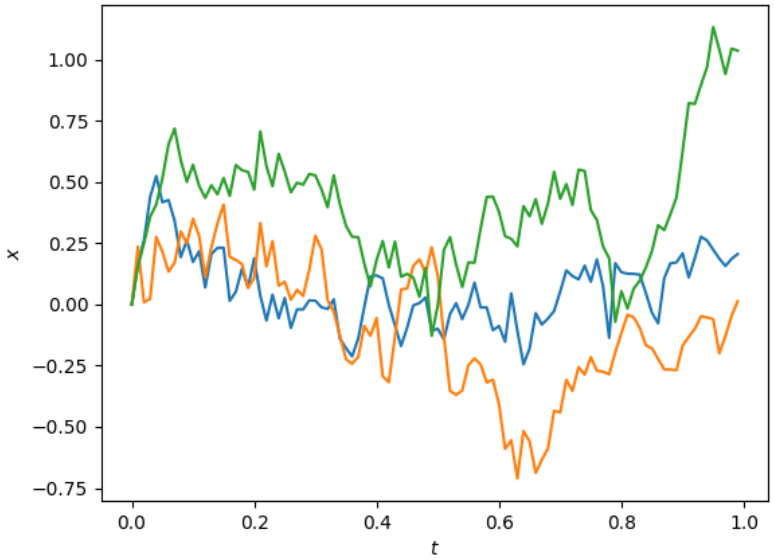


FIGURE 2.1: Three realizations of a one dimensional Wiener process ($\sigma = \sqrt{0.01}$). Here one can observe the continuity, non-differentiability of the Wiener process.

- $\forall s, t \in [0, T]$ with $s < t$ the increments $W_t - W_s$ are homogeneous and independent,
- the increments are Gaussian-distributed with mean 0 and standard deviation $\sigma = \sqrt{|t-s|}$, ergo $W_t - W_s \propto \mathcal{N}(0, |t-s|), \forall s, t \in [0, T]$,
- $W_0 = 0$,
- W_t is a continuous process,

In 1905, Einstein explained Brownian motion, what the Wiener process is also sometimes called, with the collision of many fluid molecules with the much larger pollen. The paths of the Wiener process are continuous but nowhere differentiable. This means there doesn't exist a velocity in a conventional sense [DEK60].

2.1.1 The Markov and Martingale Properties

The Wiener Process has two important properties that will come up time and again in the description of other stochastic processes. To describe them on needs to understand conditional probabilities. Conditional probabilities give the probability given a number of past events, written as

$$\mathcal{P}[X_{n+1}|X_1, X_2 \dots X_n] = p, \tag{2.4}$$

which means the probability of the event X_{n+1} happening, given that the events X_1 to X_n already happened is p. The Wiener process is a Markov process, meaning that the future of the process is not dependent on the past, only on the present state of the process, which is sometimes phrased as the process having no memory [PB13], written in terms of conditional

probabilities

$$\mathcal{P}[X_{n+1}|X_1, X_2 \dots X_n] = \mathcal{P}[X_{n+1}|X_n]. \tag{2.5}$$

This property is not limited to stochastic processes. In deterministic Newton mechanics, where a state is described by position and velocity of all particles, this property gives rise to the Laplacian demon, that all future states can be predicted by knowledge of the position and velocity of all particles at time t [LTE02]. In reinforcement learning, it is very important to know whether the problem to be solved is Markov or not due to the increased difficulty and memory needed to solve a non-Markovian problem [SB20]. In stochastic mechanics, this property simplifies the simulation and analysis of stochastic processes by a huge margin, because it it not necessary to compute the memory of the problem.
The Wiener process is also a Martingale process. If for each event A_n

$$\mathbb{E}[A_n] = 0, \qquad \mathbb{E}[A_{n+1}|A_1, \dots, A_n] = 0 \tag{2.6}$$

hold, then for a cumulative sequence

$$X_n = \sum_n A_n \tag{2.7}$$

the expectation value is

$$\mathbb{E}[X_{n+1}|X_1, \dots, X_n] = \mathbb{E}[X_{n+1}|A_1, \dots, A_n] = X_n. \tag{2.8}$$

Such a sequence is called martingale. For stochastic processes, this means that the best prediction for the next step of the process is its current value

$$\mathbb{E}[X_{n+1}|X_1, X_2 \dots X_n] = X_n. \tag{2.9}$$

Related to the martingale property are the submartingale and supermartingale properties that occur, if the best prediction is larger than the current value

$$\mathbb{E}[X_{n+1}|X_1, X_2 \dots X_n] \geq X_n. \tag{2.10}$$

or smaller than the current value

$$\mathbb{E}[X_{n+1}|X_1, X_2 \dots X_n] \leq X_n. \tag{2.11}$$

respectively. A martingale process describes a process with the mean value for the jump size of zero, ergo with equal probability of jumping “right” or “left” [PB13]. Also related is the semimartingale property, that a stochastic process X_t can be decomposed into

$$X_t = M_t + A_t \tag{2.12}$$

where M_t is a local martingale process and A_t is an adapted (also called non-anticipating), meaning that the value of X_t is known at time t and is a càdlàg function, meaning continuous from the right and having a limit to the left.

2.2 Stochastic Differential Equations

Differential equations are well suited for describing the universe. Because ordinary and partial differential equations are deterministic, we need another kind of differential equation to describe stochastic processes. One can motivate them from ordinary differential equations, for example

$$\frac{\mathrm{d}X}{\mathrm{d}t} = b(t, X(t)), \tag{2.13}$$

by multiplying by $\mathrm{d}t$,

$$\mathrm{d}X = b(t, X(t)\mathrm{d}t, \tag{2.14}$$

and adding a noise term

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}\xi, \tag{2.15}$$

where the first, deterministic term is called drift term and the second, non-deterministic term is called the diffusion term. In this thesis, only stochastic differential equations (SDEs) are treated, where the noise process $\mathrm{d}\xi$ has the properties of a Wiener process and is thus denoted as $\mathrm{d}W_t$. It is of course hypothetically possible to have stochastic differential equations with other noise processes, however this will break most if not all rules of standard stochastic calculus. Thus, the Wiener process is the only noise process used in this thesis.
To integrate SDEs

$$X_t = X_0 + \int_0^t b(t, X_t)\mathrm{d}t + \int_0^t \sigma(s, X_s)\mathrm{d}W_s \tag{2.16}$$

one can see, that the drift term can be treated with the common Lebesgue or Riemann integral. The Lebsegue integral is defined via a measure μ and a corresponding decomposition of the $f(\vec{x})$ axis c_i and corresponding sets A_i,

$$\int f(\vec{x})\mathrm{d}x^n = \sum_i c_i \mu(A_i), \tag{2.17}$$

while the Riemann-integral is defined over a decomposition of the argument axis Δx_i and the supremum or infimum of the function f resulting in so called lower and upper Darboux sums:

$$L = \sum_i \inf(f)\Delta x_i, \tag{2.18}$$

$$U = \sum_i \sup(f)\Delta x_i, \tag{2.19}$$

which equal in the limit to the Riemann integral

$$\int f(x)\mathrm{d}x = \lim_{\Delta x \to 0} \sum_i \inf(f)\Delta x_i = \lim_{\Delta x \to 0} \sum_i \sup(f)\Delta x_i. \tag{2.20}$$

The diffusion term however is a stochastic integral and has to be treated differently [Gar09].

First, we have to divide the integral into n subintervals using the partition points

$$t_0 \leq t_1 \leq \cdots \leq t_{n-1} \leq t \tag{2.21}$$

and intermediate points τ_i with

$$t_{i-1} \leq \tau_i \leq t_i, \tag{2.22}$$

similar to Riemann integrals. Now we define the stochastic integral as the limit of the partial sums

$$S_n = \sum_{i=1}^{n} \sigma(\tau_i)[W_{t_i} - W_{t_{i-1}}]. \tag{2.23}$$

However we didn't set a specific value for the intermediate points. There are two main choices one can make here. The first one is choosing the intermediate point to be the "left" edge of the interval

$$\tau_i = t_{i-1} \tag{2.24}$$

which is called the Itô integral [Gar09]

$$\int_0^t \sigma(s, X_s)\mathrm{d}W_s = \sum_{j=0}^{N-1} \sigma(t_j)(W_{t_{j+1}} - W_{t_j}) \tag{2.25}$$

for elementary functions. For general measurable functions (see Appendix A.2) we define the Itô integral using

Theorem 1 *For every measurable function $f : [0,T] \times \Omega \to \mathbb{R}^n$ exists a sequence of elementary functions ϕ_n such that*

$$\lim_{n\to\infty} \mathbb{E}\Big[\int_0^T \big(f(t) - \phi_n(t)\big)^2 \mathrm{d}t\Big] = 0 \tag{2.26}$$

almost surely (proof see [Arn73]).

Definition 1 *Let f be a measurable function $f : [0,T] \times \Omega \to \mathbb{R}^n$, then*

$$\int_0^t f(s, X_s)\mathrm{d}W_s = \textit{ms-lim}_{n\to\infty} \int_0^T \phi_n(t_j)\mathrm{d}W_t \tag{2.27}$$

is called the Itô integral with ms-lim denoting the convergence in the mean square [Kö18; Gar09; Bey23; Arn73; Øks10](see Appendix A.1).

The Itô integral has the following properties [Kö18]:

(i) it is linear,

$$\int_0^T (\alpha f(t, X_t) + \beta g(t, X_t))\mathrm{d}W_t = \alpha \int_0^T f(t, X_t)\mathrm{d}W_t + \beta \int_0^T g(t, X_t)\mathrm{d}W_t \tag{2.28}$$

with $\alpha, \beta \in \mathbb{R}$,

(ii) for $a, b \in [0, T], a < b$ holds

$$\int_0^T \mathbb{1}_{[a,b]} \mathrm{d}W_t = W_b - W_a, \tag{2.29}$$

(iii) if $0 < U < T$

$$\int_0^T f(t, X_t)\mathrm{d}W_t = \int_0^U f(t, X_t)\mathrm{d}W_t + \int_U^T f(t, X_t)\mathrm{d}W_t \tag{2.30}$$

holds for almost all X_t,

(iv) $\mathbb{E}[\int_0^T f(t, X_t)\mathrm{d}W_t] = 0$,

(v) $\mathbb{E}[\int_0^T f(t, X_t)\mathrm{d}W_t]^2 = \mathbb{E}[\int_0^T f(t, X_t)\mathrm{d}t] = \int_0^T \mathbb{E}[f(t, X_t)^2]\mathrm{d}t$,

(vi) $\int_0^T f(t, X_t)\mathrm{d}W_t$ is measurable,

(vii) The Itô integral has continuous realizations with probability 1.

With these properties one can obtain a chain rule for stochastic processes. A function of a stochastic process is itself again a stochastic process $Y_t = f(t, X_t)$ and its differential is described with the Itô formula [Bey23; Itô44; Øks10] (sometimes called Itô's lemma)

$$\begin{aligned}\mathrm{d}Y_t =& \partial_t f(t, X_t)\mathrm{d}t + \partial_{X_t} f(t, X_t)\mathrm{d}X_t + \frac{1}{2}\mathrm{d}X_t^T(\partial_{X_t}^2 f(t, X_t))\mathrm{d}X_t && (2.31)\\ =& [\partial_t f(t, X_t) + b(t, X_t)\partial_{X_t} f(t, X_t) && (2.32)\\ &+ \frac{1}{2}\mathrm{Tr}[\sigma^T(t, X_t)(\partial_{X_t}^2 f(t, X_t))\sigma(t, X_t)]]\mathrm{d}t + \mathrm{d}W_t \partial_{X_t} f(t, X_t).\end{aligned}$$

The other common choice of intermediate point is using the middle point

$$\tau_i = \frac{t_{i-1} + t_i}{2}. \tag{2.33}$$

This yields the Stratonov-integrals (see Appendix A.3) which are indicated with a circle, like

$$\int_0^T f(t, X_t) \circ \mathrm{d}W_t \tag{2.34}$$

for example. Itô integrals are more easy to solve numerically, whilst Stratonov integrals are better for stochastic calculus, especially on manifolds, because they preserve the standard rules of calculus [Bey23]. In this thesis we will focus on Itô integrals. Stochastic processes who fulfill the SDE

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t \tag{2.35}$$

are called Itô diffusion processes. All Itô diffusion processes have the semimartingale property.

2.3 Numerical Methods for Solving Stochastic Differential Equations

There are three numerical integration schemes for solving SDEs directly. The first one is similar to the Euler method used for ordinary differential equations,

$$X_{t+\Delta t} = X_t + b(t, X_t)\Delta t + \sigma(t, X_t)\Delta W_t. \tag{2.36}$$

It differs from the ordinary Euler scheme by a third term and is called Euler-Maruyama scheme. This scheme has a strong convergence order of 1 [Klo02]. To get a higher order of convergence one can use the Heun scheme, sometimes called the improved Euler scheme. Here we take an Euler step as a predictor step and then use a corrector resulting in

$$X_e = X_t + b(t, X_t)\Delta t + \sigma(t, X_t)\Delta W_t \tag{2.37}$$

$$X_{t+\Delta t} = X_t + \frac{1}{2}(b(t, X_t) + b(t + \Delta t, X_e))\Delta t + \frac{1}{2}(\sigma(t, X_t) + \sigma(t + \Delta t, X_e))\Delta W_t \tag{2.38}$$

where ΔW_t is the same in both steps. This scheme has a strong convergence order of 1 [Klo02]. The third method is the Milstein scheme (see Appendix A.5), a one-step scheme which adds a term to the Euler scheme. This however doesn't lead to a higher convergence order compared to the Euler-Maruyama scheme for constant $\sigma(t, X_t)$ and is thus not used in this thesis [Sch99]. Both methods are written here for the one-dimensional case but can easily be expanded to higher dimensions.
There exist higher-order Runge-Kutta schemes for the solution of stochastic differential equations [RI10; R09]. However they become increasingly complicated and require substantial computational effort. Also they don't necessarily lead to a higher convergence order (see Appendix (A.4) [KP11; Haa21]).

2.4 The Fokker-Planck Equation

As stated in (2.1), stochastic processes are related to a probability measure often described by a probability distribution $\rho(t, x)$. This probability distribution usually changes over time. For Itô diffusion processes (2.35), this time evolution is described by a partial differential equation, the Fokker-Planck equation

$$\partial_t \rho(t, x) = -\partial_{x_i}(b_i(t, x)\rho(t, x)) + \frac{1}{2}\partial_{x_i, x_j}(\sigma_{ik}(t, x)\sigma_{jk}(t, x)\rho(t, x)). \tag{2.39}$$

given here for the multidimensional case. One can derive this equation by using the Itô formula [Bey23; Gar85](see Appendix A.6). This means that one can also solve a SDE by solving its corresponding partial differential equation.
If one writes the Wiener process in the SDE as

$$\mathrm{d}x = \mathrm{d}W_t \tag{2.40}$$

one sees, that in the notation of the Itô diffusion, eq. (2.35), that $b(t, X_t) = 0$ and $\sigma(t, X_t) = 1$. If one inserts this into the Fokker-Planck equation, it gives the one-dimensional diffusion

equation

$$\partial_t \rho(t,x) = \frac{1}{2}\partial_x^2 \rho(t,x), \tag{2.41}$$

with $\frac{1}{2}$ as the thermal diffusivity and the boundary conditions

$$\lim_{x \to \pm\infty} \rho(x,t) = 0. \tag{2.42}$$

2.5 Example: The Langevin Equation

Another way of describing a large particle with mass M in a fluid consisting of much smaller molecules is the Langevin equation

$$\mathrm{d}v(t) = -\gamma v(t)\mathrm{d}t + \frac{\sigma}{M}\mathrm{d}W(t) \tag{2.43}$$

with its coupled SDE for the position

$$\mathrm{d}x(t) = v\mathrm{d}t \tag{2.44}$$

(γ viscosity, σ strength of the Wiener process dependent on the temperature). We will simulate this model as an example for the concepts from this chapter.
To gather some sample paths, we use the Heun scheme, leading to the numerical steps

$$V_e = V_t - \gamma V_t \Delta t + \frac{\sigma}{M}\Delta W_t, \tag{2.45}$$

$$V_{t+\Delta t} = V_t - \frac{\gamma}{2}(V_t + V_e)\Delta t + \frac{\sigma}{M}\Delta W_t, \tag{2.46}$$

$$X_e = X_t + V_e \Delta t, \tag{2.47}$$

$$X_{t+\Delta t} = X_t + \frac{1}{2}(V_{t+\Delta t} + V_e)\Delta t, \tag{2.48}$$

with the initial values $X_0 = 0, V_0 = 0$. For the sake of simplicity, we set $\gamma = \sigma = M = 1$ for this simulation and $\Delta t = 0.01$ and limit ourselves to one-dimensional motion. In Fig. 2.2 one can see large, quick changes in velocity but only small/slow changes in position. The probability distribution of the velocity is described by the corresponding Fokker-Planck equation

$$\partial_t \rho(t,v) = \gamma \partial_v (v\rho(t,v)) + \frac{\sigma^2}{M^2}\frac{1}{2}\partial_v^2 \rho(t,v). \tag{2.49}$$

Equation (2.43) has similarities with another stochastic process, namely

$$\mathrm{d}x = -\alpha x \mathrm{d}t + \sigma \mathrm{d}W, \tag{2.50}$$

the Ornstein-Uhlenbeck process. For the Ornstein-Uhlenbeck process the probability distribution is the Gaußian distribution

$$\rho(x,t) = \mathcal{N}\left(X_{t=0}e^{-\alpha t}, \quad \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha t}\right)\right), \tag{2.51}$$

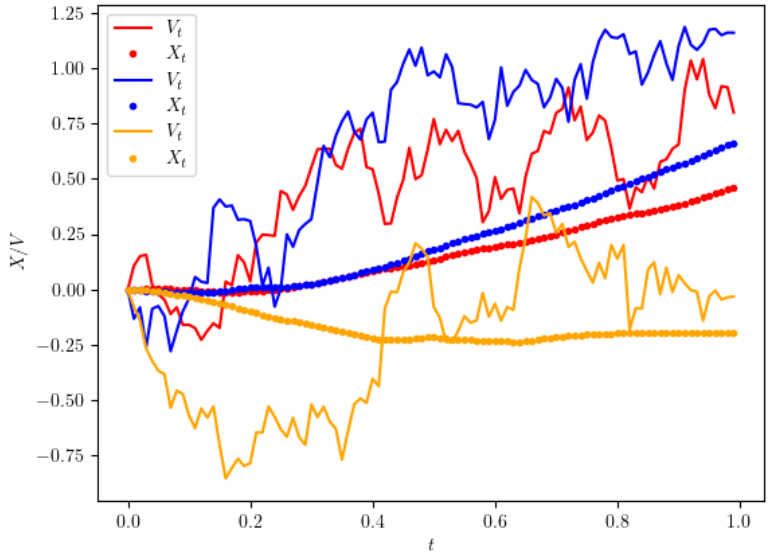


FIGURE 2.2: Three sample paths of the velocity and its corresponding position of the one dimensional Langevin equation with $X_0 = 0, V_0 = 0, \gamma = \sigma = M = 1$ integrated via the Heun scheme.

which means that this is also the probability distribution for the velocity of the Langevin equation

$$\rho(v,t) = \mathcal{N}\left(V_{t=0}e^{-\gamma t}, \frac{\sigma^2}{2M^2\gamma}\left(1 - e^{-2\alpha t}\right)\right). \tag{2.52}$$

Chapter 3

Nelson's Stochastic Interpretation of Quantum Mechanics

After Schrödinger argued the equation, now named after him, into existence in 1926 [Sch26b; Sch26c; Sch26d; Sch26e], it quickly, together with the matrix formulation, which is equivalent to the wave mechanics [Sch26f], how it was called at the time, became the only way to look at quantum mechanics. However, Schrödinger's aim was to find the quantum mechanical equivalence to the classical Hamilton-Jacobi formalism. So why should not there be other formalisms for quantum mechanics? In this chapter, I will give a short recapitulation of the formalisms of classical mechanics and how they are interconnected. Then the stochastic mechanics approach to quantum mechanics will be introduced with its connection to the Schrödinger equation. Eventually, advantages and disadvantages and connection to other quantum mechanical formalisms will be disclosed.

3.1 Basics of Classical Mechanics

Isaac Newton developed the foundation for theoretical classical mechanics with Newton's three laws of motion. Especially, the second Newton equation

$$m\ddot{\vec{x}} = \vec{F} \tag{3.1}$$

transforms the search for the trajectory of a particle into solving an initial value problem. For conservative forces one can define a potential with

$$\vec{F} = -\vec{\nabla}V. \tag{3.2}$$

Rowan Hamilton was looking for an underlying principle that determined the specific trajectory of particles and derived the Hamilton principle

$$\delta S = \delta \int_0^T L(x_1, \ldots, x_N, \dot{x_1}, \ldots, \dot{x_N}, t)\mathrm{d}t = 0, \tag{3.3}$$

which means the particles, take the trajectory which minimizes the action functional. From variational calculus (see Appendix B.1), we know that the functional is extremized, if the

trajectories obey the Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}_i}\right) - \frac{\partial L}{\partial x_i} = 0 \tag{3.4}$$

for each coordinate x_i. In general one uses here q_i for the generalized coordinates to indicate that we can also describe motion on non-Euclidean manifolds. This notation is omitted here. This partial differential equation is the core of Lagrangian mechanics. Until now, we didn't specify how the Lagrangian L should look like. To recover Newtonian mechanics, we set $L = T - V$ with T being the kinetic energy, so for a single particle on a one dimensional Euclidean manifold

$$L(x, \dot{x}) = \frac{m}{2}\dot{x}^2 - V(x), \tag{3.5}$$

and if we put this Lagrangian into the Euler-Lagrange equation

$$0 = m\ddot{x} + \partial_x V(x) \tag{3.6}$$

we recover Newton's second law (3.1). One can go now from configuration space to phase space by performing a Legendre transformation. First we need the conjugated generalized momentum

$$p_i = \frac{\partial L}{\partial \dot{x}_i} \tag{3.7}$$

to arrive at the transformation

$$H(x_1, \ldots, x_N, p_1, \ldots, p_N, t) = \sum_{i=1}^{N} p_i \dot{x}_i - L(x_1, \ldots, x_N, \dot{x}_1, \ldots, \dot{x}_N, t) \tag{3.8}$$

where H is called the Hamiltonian or the Hamilton function. Taking the total differential of H yields

$$\begin{aligned}
\mathrm{d}H &= \sum_{i=1}^{N}(\mathrm{d}p_i\dot{x}_i - \mathrm{d}x_i p_i) - \sum_{i=1}^{N}(\partial_{x_i} L\mathrm{d}x_i + \partial_{\dot{x}_i} L\mathrm{d}\dot{x}_i) - \partial_t L\mathrm{d}t && (3.9)\\
&= \sum_{i=1}^{N}(\dot{x}_i\mathrm{d}p_i - \partial L_{x_i}\mathrm{d}x_i) - \partial_t L\mathrm{d}t && (3.10)\\
&= \sum_{i=1}^{N}(\dot{x}_i\mathrm{d}p_i - \dot{p}_i\mathrm{d}x_i) - \partial_t L\mathrm{d}t. && (3.11)
\end{aligned}$$

Comparing this expression with the definition of the total differential

$$\mathrm{d}H = \sum_{i=1}^{N}(\partial_{p_i} H\mathrm{d}p_i + \partial_{x_i} H\mathrm{d}x_i) + \partial_t H\mathrm{d}t \tag{3.12}$$

we deduce the Hamilton equations

$$\dot{x}_i = \frac{\partial H}{\partial p_i}, \tag{3.13}$$

$$\dot{p}_i = -\frac{\partial H}{\partial x_i}, \tag{3.14}$$

also called the canonical equations, which are the core of Hamiltonian mechanics. To recover Newtonian mechanics, we set the Hamilton function

$$H = T + V, \tag{3.15}$$

so for the example above

$$H(x,p) = \frac{p^2}{2m} + V(x). \tag{3.16}$$

The Hamilton equations then become

$$\dot{x} = \frac{p}{m}, \tag{3.17}$$

$$\dot{p} = -\partial_x V(x). \tag{3.18}$$

Performing the total derivative of (3.17) with respect to time and inserting it into equation (3.18), eliminating p, we get

$$m\ddot{x} = F, \tag{3.19}$$

Newton's second law.
In phase space, one can not simply change variables. If one wants to change variables and still keep the original canonical equations one needs a canonical transformation. Now, we want to find a canonical transformation that makes our set of variables $\tilde{x}_i, \tilde{p}_i$ cyclic (meaning $\dot{\tilde{p}}_i = 0, \forall i = 1, \ldots, N$). For this we transform the Hamiltonian using a generating function of the second kind $F_2(x_i, \tilde{p}_i, t)$

$$\tilde{H} = H + \frac{\partial F_2}{\partial t} = 0. \tag{3.20}$$

Using

$$p_i = \frac{\partial F_2}{\partial x_i}, \tag{3.21}$$

$$\tilde{x}_i = \frac{\partial F_2}{\partial \tilde{p}_i}, \tag{3.22}$$

we obtain the Hamilton-Jacobi equation

$$H\left(x_1, \ldots, x_N, \frac{\partial F_2}{\partial x_1}, \ldots, \frac{\partial F_2}{\partial x_n}, t\right) + \frac{\partial F_2}{\partial t} = 0, \tag{3.23}$$

the core of Hamilton-Jacobi mechanics [Nol14].
To show the physical meaning of the generating function $F_2(x_i, \tilde{p}_i, t)$ we take its total time derivative

$$\frac{\mathrm{d}F_2}{\mathrm{d}t} = \sum_{i=1}^{N} \left(\frac{\partial F_2}{\partial x_i}\dot{x}_i + \frac{\partial F_2}{\partial \tilde{p}_i}\dot{\tilde{p}}_i \right) + \frac{\partial F_2}{\partial t} \tag{3.24}$$

and use equations (3.20), (3.21) and (3.22) together with $\dot{\tilde{p}}_i = 0$ to find

$$\frac{\mathrm{d}F_2}{\mathrm{d}t} = \sum_{i=1}^{N} p_i \dot{x}_i - H \tag{3.25}$$

where the right hand side is the inverse Legendre transformation from phase space to configuration space, ergo from the Hamilton to the Lagrange function

$$\frac{\mathrm{d}F_2}{\mathrm{d}t} = L. \tag{3.26}$$

Integrating this function

$$F_2 = \int_0^T L \mathrm{d}t, \tag{3.27}$$

we find that F_2 is the action functional S from the Hamilton principle (3.3) [Nol14]. To recover Newtonian mechanics, the Hamilton-Jacobi equation has to have the form

$$\sum_{i=1}^{N} \frac{1}{2m} \frac{\partial S}{\partial x_i} + V(x_1, \dots x_N) + \frac{\partial S}{\partial t} = 0, \tag{3.28}$$

taken from the form of the Hamiltonian (3.15).
This derivation shows that there are at least four, more or less equivalent (see Appendix B.2) [CA21], mathematical descriptions of classical mechanics. Schrödinger with his equation wanted to recover Hamilton-Jacobi mechanics for quantum mechanics.

3.2 Standard Formulation of Quantum Mechanics

Classical physics failed to explain certain phenomena at the brink of the 20th century. The most prominent is probably the photoelectric effect. When monochromatic light is send on a metal surface, electrons are emitted if the frequency is high enough. The "classically" problematic part is that for lower frequencies one could not simply increase the intensity of the light to get electrons to be emitted. This is of course a contradiction to the wave interpretation of light. This problem was solved by Planck [Pla01] and Einstein [Ein05] by introducing photons, quantized energy particles with the energy

$$E = hf. \tag{3.29}$$

This contradicted the classical understanding of light as an electromagnetic wave (Maxwell [Max65]).
Another problem was the structure of the atom. Ernest Rutherford found in 1911 that the atom was build of a heavy nucleus and electrons rotating around that nucleus [Rut12]. But these electrons, being accelerated charges, should emit electromagnetic waves, loose energy, and collapse into the nucleus, resulting in unstable atoms.
Another contradiction to classical mechanics was observed, when the double slit experiment was repeated with particles and interference was observed, a feature previously exclusively associated with waves, but not with particles. A solution was found in the Schrödinger

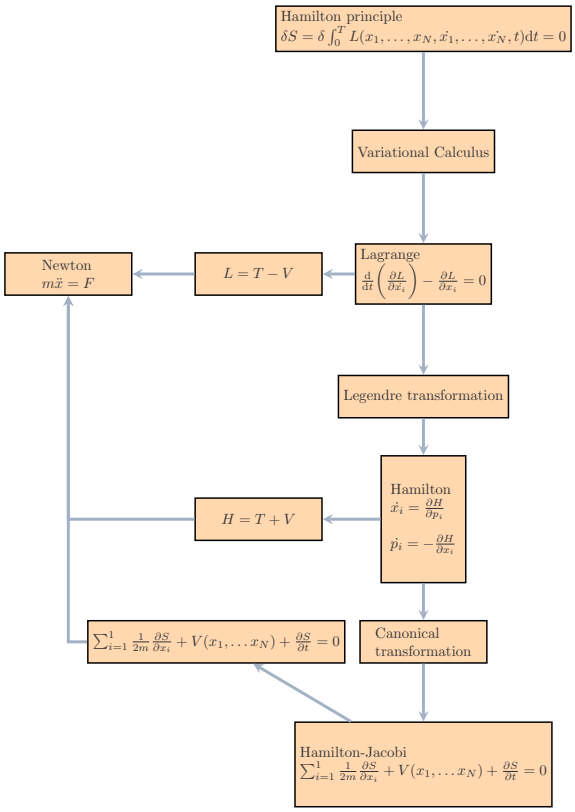


FIGURE 3.1: Classical formalisms and their connections to each other.

equation

$$i\hbar\partial_t \Psi = \hat{H}\Psi, \tag{3.30}$$

or

$$\hat{H}\Psi = E\Psi \tag{3.31}$$

in its stationary form by using complex arguments to generalize the Hamilton-Jacobi-equation for the quantum realm. The complex function $\Psi(x,t)$ is called the wave function (despite the Schrödinger equation not being a wave equation but a heat equation with complex coefficients). Schrödinger didn't derive the Schrödinger equation but argued it into existence [Sch26b; Sch26c; Sch26d; Sch26e]. He started with the ansatz of matter waves, first introduced by de Broglie in his dissertation in 1924 [Bro24].

Using the Hamilton-Jacobi formalism for conservative systems one can write the action

$$S(\vec{r},\vec{\tilde{p}},t) = W(\vec{r},\vec{\tilde{p}}) - Et. \tag{3.32}$$

With the cyclic nature of $\vec{\tilde{p}}$, W = const. defines an area in configuration space, that the S = const.-areas, the action waves, move through. Using

$$0 = \mathrm{d}S = \nabla_r W \mathrm{d}\vec{r} - E\mathrm{d}t \tag{3.33}$$

one can define a wave and phase speed $\vec{u}$ by

$$\nabla_r W \vec{u} = E, \tag{3.34}$$

which is orthogonal to the action wave fronts. The momentum of the particle is

$$\vec{p} = \nabla_r W = m\vec{v} \tag{3.35}$$

and because the particle speed must be parallel to the wave speed, one can write

$$u = \frac{E}{p} = \frac{E}{mv} \rightarrow uv = \text{const.} \tag{3.36}$$

Thus we can derive the classical wave function by inserting this energy formula into the Hamilton-jacobi equation

$$H(x_1,\ldots,x_N,\frac{\partial S}{\partial x_1},\ldots,\frac{\partial S}{\partial x_N}) + \frac{\partial S}{\partial t} = 0 \tag{3.37}$$

with (3.32) and get the equation

$$(\nabla_r S)^2 = \frac{1}{u^2}\left(\frac{\partial S}{\partial t}\right)^2, \tag{3.38}$$

the wave equation for action waves. This gives rise of the so called wave-particle dualism. From this, one can search for similarities between light waves and matter waves. Starting with the wave function for the scalar potential known from classical electrodynamics

$$\nabla^2\phi - \frac{n^2}{c^2}\frac{\partial^2\phi}{\partial t^2} = 0, \tag{3.39}$$

(c the speed of light in vacuum, n the refractive index). Per analogy we find that

$$u = \frac{c}{n} \tag{3.40}$$

and with n = const. the plane wave

$$\phi(\vec{r},t) = \phi_0 e^{i(\vec{k}\vec{r}-\omega t)} \tag{3.41}$$

solves this equation with the dispersion relation

$$k = \omega\frac{n}{c} = \frac{\omega}{u} = \frac{2\pi}{\lambda}. \tag{3.42}$$

If $n = n(\vec{r})$, we take the ansatz

$$\phi(\vec{r}, t) = \phi_0(\vec{r}) e^{i\frac{k}{n}(L\vec{r}-ct)}, \tag{3.43}$$

which leads to the eikonal equation

$$(\nabla L(\vec{r}))^2 = n = \frac{c^2}{u^2}. \tag{3.44}$$

With $L = \text{const.}$ we analogously get the wave fronts similar to the above. The corresponding trajectories are the light beams known from optics, the special case of wave optics. Thus it is alluring to postulate an analog wave mechanics for matter for which classical mechanics is a special case. Thus from comparison, the phase factor from (3.43) for the action waves is $S = W - Et$, which means

$$E \propto \frac{k}{n}c = ku \propto \frac{u}{\lambda} = f, \tag{3.45}$$

so we obtain (3.29), Planck's ansatz for the energy of photons. With

$$\lambda = \frac{u}{f} = \frac{E}{pf} = \frac{h}{p} \tag{3.46}$$

we obtain for the momentum

$$p = \frac{h}{\lambda} \tag{3.47}$$

where λ is the de Broglie wavelength. Using (3.43) to eliminate the time-derivative in the wave equation, we obtain

$$\Delta_r \phi + \frac{4\pi^2}{\lambda^2} = 0, \tag{3.48}$$

and with

$$\frac{4\pi^2}{\lambda^2} = \frac{p^2}{\hbar^2} = \frac{2m}{\hbar^2}(E - V) \tag{3.49}$$

we obtain the stationary Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\Delta + V \right)\Psi = E\Psi \tag{3.50}$$

with Ψ instead of ϕ to differentiate this from the light waves. The part in parenthesis is defined as the Hamilton operator

$$\hat{H} = -\frac{\hbar^2}{2m}\Delta + V(\vec{r}). \tag{3.51}$$

Using the analogy from the wave equation and (3.43), we get

$$E \to i\hbar\frac{\partial}{\partial t} \tag{3.52}$$

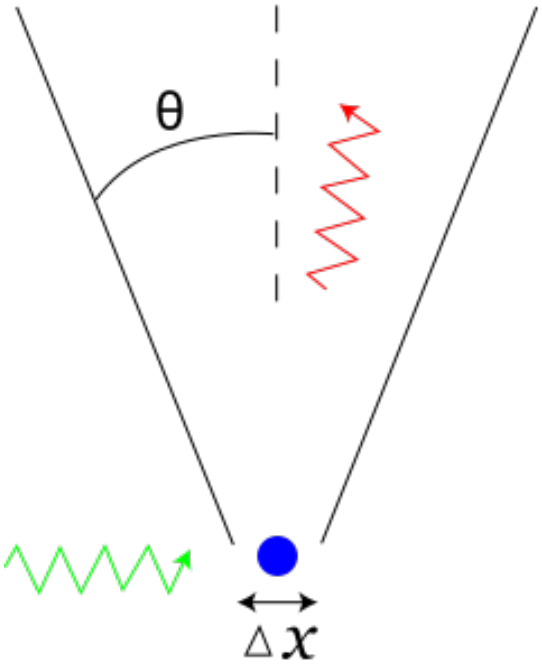


FIGURE 3.2: Scheme of the gamma ray microscope thought experiment (taken from Wikimedia Commons [Com20]).

and

$$\hat{H}\Psi(\vec{r},t) = i\hbar\frac{\partial}{\partial t}\Psi(\vec{r},t), \tag{3.53}$$

the time-dependent Schrödinger equation [Nol13; Sch26b; Sch26c; Sch26d; Sch26e].
The interpretation of this equation was a matter of debate and the interpretation of the wave function not clear. The solution was the Born rule

$$\rho(x) = |\Psi|^2, \tag{3.54}$$

which connected the wave function to the probability density and gave quantum mechanics a statistical interpretation [Bor26]. This of course imposes the normalization condition

$$1 = \int |\Psi|^2 \mathrm{d}x \tag{3.55}$$

for the wave function. In this interpretation Ψ is not something that exists in the real world but is a mathematical tool to describe the statistical behavior of particles.
The wave function Ψ is part of the separable Hilbert space $\mathcal{H}$ of square-integrable complex-valued functions. Observables are represented by self-adjoint operators,

$$\int \Psi^\dagger(\hat{A}\Psi)\mathrm{d}\vec{r} = \int (\hat{A}^\dagger\Psi^\dagger)\Psi\mathrm{d}\vec{r}, \tag{3.56}$$

with a spectrum of real eigenvalues.
Another important property is the interdependence of measurements. Heisenberg illustrated the problem in his γ-microscope thought experiment [Hei27]. In contrast to a classical setting, the influence of a measurement of an observable on a particle is not negligible. In his thought

experiment, Heisenberg wanted to measure position and momentum of a particle with a single photon, hence the name γ-microscope. From optics, we know that the uncertainty of the position measurement is proportional to the wavelength λ. In the moment, at which the photon hits the electron, is reflected or scattered, the momentum of the electron is changed. The change is the larger, the smaller the wavelength, ergo the more certain we are of the position, the more uncertain we are of the momentum [Hei27]. In the same year E.H. Kennard derived

$$\Delta\vec{p}\Delta\vec{x} \geq \frac{h}{2\pi} = \hbar, \tag{3.57}$$

the Heisenberg uncertainty relation in its most common notation [EH27].
To find out, if one can measure two different observables interchangeably, one applies the operators in both permutations and looks at the difference. If this difference is zero, one can measure the corresponding observables interchangeably; if not, one can not measure both observable arbitrarily accurate independently from one another, which is not the case for the position and the momentum operator:

$$\hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar = [\hat{x}, \hat{p}]. \tag{3.58}$$

$[\cdot, \cdot]$ is called commutator. Here it should be mentioned, that there are two schools with a slight difference in interpretation. The common way to view the wave function is a way to describe the behavior of a particle. The other, less common, way first proposed by Ballentine [BAL70] sees the wave function as a statistical ensemble of many measurements. In the second view, it should be theoretically possible to break uncertainty relations in single experiments. This however has not been checked experimentally yet.

When the nature of the quantum realm was revealed and the standard mathematical framework was developed in the beginning of the 20th century, the interpretation of the quite unusual findings were intensively discussed. The most commonly taught is the Copenhagen interpretation. Here, the wave function is not interpreted as real but as a description of our knowledge of the problem to solve. During the measurement with a classical apparatus, there is an instantaneous collapse of said wave function, the wave nature of the particle disappears and the particle appears to us as such. The duality of wave and particle nature are coupled in the concept of complementarity, which says that the experiment determines the nature of the quantum particle - wave or particle.
This interpretation has some shortcomings. One is the anti-realism of the complementarity concept, which erases the concept of an objective observer, a quite important concept of modern science [Keu07; Fra19]. Another one is the arbitrary separation of the classical and the quantum world. And last but not least, the interpretation of the wave function as something purely mathematical displeases many physicists, who want to give physical meaning to this integral part of conventional quantum mechanics. These inconsistencies are the the reason why the Copenhagen interpretation is sometimes derogatorily called the "Shut up and calculate" interpretation after the famous quote by David Mermin:

> "If I were forced to sum up in one sentence what the Copenhagen interpretation says to me, it would be: "Shut up and Calculate! But I won't shut up." [DM89].

More in regards to the different interpretations of quantum mechanics will be said in a later

section of this thesis.

3.3 Nelson Mechanics

Edward Nelson formulated the stochastic description of quantum systems in 1966 [Nel66]. The main idea is that quantum systems can not be viewed as closed systems. Nelson started from three assumptions

- The particle is subject to a mean acceleration $\mathbb{E}[a]$ given by Newton's second law

$$F = m\mathbb{E}[a]. \tag{3.59}$$

- The particle is undergoing Brownian motion by being coupled to a stochastic background field, which gives the diffusion coefficient $\sigma^2 = \frac{\hbar}{m}$. $\sigma \propto m^{-1}$, because for heavier objects the Brownian motion is irrelevant and we recover Newton mechanics. $\sigma \propto \hbar$ will be needed later to recover the Schrödinger equation.
- The diffusion process is conservative, ergo non-dissipative, and thus can be described by a time-reversible stochastic process.

The derivation starts with the following n-dimensional stochastic differential equation of motion

$$\mathrm{d}X(t) = C_f(X(t), t)\mathrm{d}t + \sigma \mathrm{d}W_f(t) \tag{3.60}$$

with $C_f(X(t), t)$ being the forward drift and $\mathrm{d}W_f(t)$ the forward Wiener process. Most stochastic processes are dissipative. In quantum mechanics, we need an on average constant energy. According to the Noether theorem, energy conservation either in the mean or exactly results from time reversibility. Thus we need a corresponding SDE backwards in time,

$$\mathrm{d}X(t) = C_b(X(t), t)\mathrm{d}t + \sigma \mathrm{d}W_b(t) \tag{3.61}$$

with C_b being the backward drift, $W_b(t)$ being the backward Wiener process.
In the next step we derive the form of the drift terms, which have to have the units of a velocity. As described in the previous chapter, the stochastic processes are nowhere differentiable. Thus, using the Markov property of the Wiener process and the Brownian motion, we introduce alternative mean forward and backwards derivatives

$$D_f X(t) := \lim_{\Delta t \to 0} \mathbb{E}\left[\frac{X(t+\Delta t) - X(t)}{\Delta t}\middle| X(t)\right] \tag{3.62}$$

$$D_b X(t) := \lim_{\Delta t \to 0} \mathbb{E}\left[\frac{X(t) - X(t+\Delta t)}{\Delta t}\middle| X(t)\right] \tag{3.63}$$

with the forwards and backwards drift given by

$$D_f X(t) = C_f(X_t, t) \tag{3.64}$$

and

$$D_b X(t) = C_b(X_t, t) \tag{3.65}$$

and with the acceleration of the particle being

$$\frac{F}{m} = \mathbb{E}[a(t)] := \frac{1}{2}(D_b D_f + D_f D_b)X(t). \tag{3.66}$$

From the corresponding forward and backward Fokker-Planck equations

$$\partial_t \rho(x,t) = -\vec{\nabla} \cdot (C_f(x,t)\rho(x,t)) + \sigma\Delta\rho, \tag{3.67}$$
$$\partial_t \rho(x,t) = -\vec{\nabla} \cdot (C_b(x,t)\rho(x,t)) - \sigma\Delta\rho, \tag{3.68}$$

we derive by adding the two equations

$$\partial_t \rho(x,t) = -\frac{1}{2}\vec{\nabla} \cdot ((C_f(x,t) + C_b(x,t))\rho(x,t)) \tag{3.69}$$

the continuity equation

$$\partial_t \rho(x,t) = -\vec{\nabla} \cdot (v(x,t)\rho(x,t)) \tag{3.70}$$

with the current velocity

$$v(x,t) := \frac{1}{2}(C_f(x,t) + C_b(x,t)). \tag{3.71}$$

Subtracting the two Fokker-Planck equations, we get

$$\vec{\nabla} \cdot (C_f\rho - C_b\rho) = 2\sigma\vec{\nabla}\vec{\nabla}\rho \tag{3.72}$$

and after integration we obtain

$$C_f\rho - C_b\rho = 2\sigma\vec{\nabla}\rho, \tag{3.73}$$

with another velocity field

$$u := \frac{1}{2}(C_f - C_b), \tag{3.74}$$

called the osmotic velocity, which is connected to the probability density by

$$u = \frac{\sigma^2}{2}\frac{\nabla\rho}{\rho}. \tag{3.75}$$

Comparing the sums of u and v

$$v + u = \frac{1}{2}C_f + \frac{1}{2}C_b + \frac{1}{2}C_f - \frac{1}{2}C_b = C_f \tag{3.76}$$
$$v - u = \frac{1}{2}C_f + \frac{1}{2}C_b - \frac{1}{2}C_f + \frac{1}{2}C_b = C_b \tag{3.77}$$

one can express the forward and backwards equations of motion of the particles as

$$\mathrm{d}X(t) = \big(v(X(t),t) + u(X(t),t)\big)\mathrm{d}t + \sigma\mathrm{d}W_f(t), \tag{3.78}$$
$$\mathrm{d}X(t) = \big(v(X(t),t) - u(X(t),t)\big)\mathrm{d}t + \sigma\mathrm{d}W_b(t). \tag{3.79}$$

These are coupled forward-backward stochastic differential equations (FBSDEs) that are connected by the continuity equation (3.70) and describe a Markov process.

3.3.1 Connection to the Schrödinger Equation

The next step is to show that the equations of Nelson mechanics are equivalent to the Schrödinger equation. From the formulations from the previous chapter, we will derive the Madelung equations, which are equivalent to the Schrödinger equation for node-free or ground states.
Starting with an arbitrary, smooth function $f(X,t)$ and the Itô formula

$$df = \left(\frac{\partial f}{\partial t} + C_{f/b}\frac{\partial f}{\partial x} \pm \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}\right)\mathrm{d}t + \sigma\frac{\partial f}{\partial x}\mathrm{d}W_{f/b} \tag{3.80}$$

and the zero mean of the Wiener process yield the mean derivatives

$$D_{f/b}f(X_t,t) = [\partial_t + C_{f/b}\nabla \pm \frac{\sigma^2}{2}\nabla^2]f(X_t,t). \tag{3.81}$$

Using the second stochastic derivative or the mean acceleration (3.66) we get

$$\mathbb{E}[a] = \frac{1}{2}(D_f C_b + D_b C_f) \tag{3.82}$$

$$= \frac{1}{2}\Big[\partial_t + C_f\nabla + \frac{1}{2}\sigma^2\Delta\Big]C_b + \frac{1}{2}\Big[\partial_t + C_b\nabla - \frac{1}{2}\sigma^2\Delta\Big]C_f \tag{3.83}$$

$$= \partial_t v + (v\nabla)v - (u\nabla)u - \frac{1}{2}\sigma^2\Delta u \tag{3.84}$$

and using $F = m\mathbb{E}[a] = -\nabla V$

$$\partial_t v = -\frac{1}{m}\nabla V + (v\nabla)v + (u\nabla)u + \frac{1}{2}\sigma^2\Delta u \tag{3.85}$$

The partial time derivative of the osmotic velocity can be written

$$\partial_t u = \frac{\sigma^2}{2}\partial_t\left(\frac{\nabla\rho}{\rho}\right) \tag{3.86}$$

$$= \frac{\sigma^2}{2}\nabla\frac{\partial_t\rho}{\rho} \tag{3.87}$$

and with the continuity equation

$$\partial_t u = -\frac{\sigma^2}{2}\nabla\frac{\nabla(v\rho)}{\rho} \tag{3.88}$$

$$= -\nabla(\frac{\sigma^2}{2}\nabla v - vu) \tag{3.89}$$

we have

$$\partial_t u = -\frac{1}{2}\sigma^2\Delta v - \Delta(uv). \tag{3.90}$$

For the ground state, $v = 0$, the osmotic velocity is constant in time (but not in space). These coupled non-linear partial differential equations (3.85) and (3.90) describe the entire stochastic process but are difficult to solve.
These two partial differential equations are now be derived from the Schrödinger equation

from the two assumptions

$$u = \frac{1}{2}\nabla \ln \rho(x,t) = \sigma^2 \nabla R(x,t) \tag{3.91}$$

and

$$v(x,t) = \frac{1}{m}\nabla S(x,t) \tag{3.92}$$

stemming from the wave function

$$\Psi = e^{R+\frac{i}{\hbar}S}. \tag{3.93}$$

With these the Schrödinger equation reads

$$i\hbar(\partial_t R + \frac{i}{\hbar}\partial_t S)\Psi = -\frac{\hbar^2}{2m}(\Delta(R+\frac{i}{\hbar}S) + \nabla(R+\frac{i}{\hbar}S)^2)\Psi + V\Psi \tag{3.94}$$

with after dividing by Ψ can be divided into real and imaginary parts

$$\partial_t S + V(x,t) + \frac{1}{2m}[\nabla S]^2 - \frac{\hbar^2}{2m}\left[\frac{\Delta R}{R}\right] = 0, \tag{3.95}$$

$$\partial_t R + \frac{1}{2m}R(\Delta S)^2 + \frac{1}{m}\nabla R \nabla S = 0, \tag{3.96}$$

resulting in the two Madelung equations. Using the two assumptions (3.91) and (3.92) and applying the nabla operator on both equations, we obtain two partial differential equations for the two velocity fields(3.87) and (3.85).

The Madelung equations were first published by Erwin Madelung in 1926 [Mad26]. This description is also called the hydrodynamic picture of quantum mechanics due to its resemblance to the hydrodynamic Euler equations [Red17], if one replaces R and S by ρ and v. With the wave function ansatz $\Psi = \sqrt{\rho}e^{\frac{i}{\hbar}S}$, one obtains hydrodynamic equations for the probability density and the current velocity

$$\partial_t \rho = \nabla(\rho v), \tag{3.97}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}v = \partial_t v + (v\nabla)v. \tag{3.98}$$

Madelung wanted to derive a more graphic description of quantum mechanics. In his initial description, a particle is not point-like but rather a continuous density distribution which behaves like a fluid [Mad26]. Nowadays however, the interpretation of ρ as the probability density is more common [Red17]. Since the Madelung equations are non-linear partial differential equations, they are difficult to solve and are thus not often used in praxis.
Now, we know how stochastic mechanics describes quantum systems. But to actually use it, we have to know the velocity fields u and v. To obtain these from stochastic mechanics and not from the Schrödinger equation, one has to derive the quantum Hamilton equations.

3.4 Derivation of the Quantum Hamilton Equations of Motion

If the velocities u and v are known, equations (3.78) and (3.79) are enough to get sample paths for the problem at hand. But if they are not known, one has to develop the differential equations with equivalence to the Hamilton equations of motion in classical mechanics, the quantum Hamilton equations. The equations (3.78) and (3.79) were derived from Newtonian mechanics. Like in classical mechanics, shown in a previous chapter, one can also look towards variational principles to get different methods of describing stochastic mechanical problems. We start with the equivalent to the least-action-principle

$$J_1[v,u] = \min_v \max_u \mathbb{E}[\int_0^T \mathcal{L}(X_t, u_t, v_t, t)\mathrm{d}t + S(x_0)]. \tag{3.99}$$

We are searching for the optimal u_t and v_t, that minimize the functional. Both u_t and v_t are feedback controls, meaning they depend on the position and time of the particle [Pav95]. We choose the Lagrangian

$$\mathcal{L} = \frac{m}{2}(v_t^2 - u_t^2) - V(t, X_t) \tag{3.100}$$

in equivalence to classical mechanics with the additional term $-\frac{m}{2}u_t^2$ and in accordance to Guerra's optimal control ansatz [GM83]. The sign of the osmotic velocity can also be chosen to be positive like in [Yas81a; Yas81b; Yas83; Yas80]. To motivate the minus sign, we can take the expectation value and use (3.91)

$$\mathbb{E}\left[-\frac{m}{2}u_t^2\right] = \int\left(\frac{\hbar}{2m}\frac{\Delta\rho(t,x)}{\rho(t,x)}\right)\rho(x,t)\mathrm{d}x = \mathbb{E}[V_Q] \tag{3.101}$$

which is the expectation value of the quantum potential from Bohmian pilot wave theory, thus the negative sign. Due to the different signs of the velocities, we have to minimize the action with regard to the osmotic velocity and maximize the action with regard to the drift velocity.
The second variational principle introduced by Pavon [Pav95] is

$$J_2[v,u] = \max_v \min_u \mathbb{E}[\int_0^T mvu\mathrm{d}t + \hbar R(x_0)] \tag{3.102}$$

an entropy production principle. The configurational entropy of the system

$$S_E(t) = -\int \rho(t,x)\ln(\rho(t,x))\mathrm{d}x \tag{3.103}$$

with

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \int -\frac{\partial\rho(x,t)}{\partial t}(\ln(\rho(x,t)) + 1)\mathrm{d}x \tag{3.104}$$

$$= \int \nabla(v(x,t)\rho(x,t))(\ln(\rho(x,t)) + 1)\mathrm{d}x - \int 2mvu\rho\mathrm{d}x \tag{3.105}$$

$$\frac{1}{2}(S(0) - S(T)) = \mathbb{E}\left[\int_0^T mvu\mathrm{d}t\right] \tag{3.106}$$

is to be maximized. Thus, in quantum mechanics there is one more variational principle than in classical mechanics. The entropy production principle disappears in classical mechanics due to $u = 0$ in the classical limit.
Now, one can combine these two variational principles into one single principle, by defining a quantum velocity from the osmotic and drift velocities in the complex plane

$$v_q = v - iu, \tag{3.107}$$

resulting in the quantum Hamilton principle

$$J[v_q] = \text{extremize}_{v_q} \mathbb{E}\Big[\int_0^T \frac{m}{2} v_q^2 - V \mathrm{d}t + \Phi_0(x_0) \Big], \tag{3.108}$$

sometimes all so called the Pavon principle. It separates into the two variational principles by taking the real and imaginary parts of the cost functional. $\Phi_0(x_0)$ is an initial cost under the constraint that $\exp\{\frac{i}{\hbar}\Phi_0(x)\}$ is sufficiently integrable (has L1 norm).
The obtain the SDE, which obeys the quantum Hamilton principle, we combine Nelsons forward (3.78) and backward equations (3.79) into the complex SDE

$$dX_t = v_q(t, X_t)\mathrm{d}t + \frac{\sigma}{2}((1-i)\mathrm{d}W_f + (1+i)\mathrm{d}W_b), \tag{3.109}$$

with the diffusion part being simplified by defining the quantum noise

$$\mathrm{d}W_q := \frac{1}{2}((1-i)\mathrm{d}W_f + (1+i)\mathrm{d}W_b). \tag{3.110}$$

We recover (3.78) and (3.79) if we again take the real and imaginary part. Like in classical mechanics, we have to find a Hamiltonian, which corresponds to the quantum Hamilton principle. For this, we need to find the correct conjugate variables. Due to the stochasticity, we get two costate variables, instead of only one in classical mechanics. From [BG10] we obtain the Hamiltonian as

$$H(t, x, v_q, p, Q) = -\frac{m}{2} v_q^2 + V(x) + p v_q - \sqrt{\frac{\hbar}{m}} \frac{1+i}{2} Q \tag{3.111}$$

with $p(t)$ and $Q(t)$ being costate processes. To obtain the extremum of the Pavon principle, we set

$$\frac{\partial H}{\partial v_q} = 0, \tag{3.112}$$

resulting in

$$p(t) = m v_q(t). \tag{3.113}$$

From [BG10] we also obtain a backwards stochastic differential equation for the costate process

$$\mathrm{d}p(t) = -\partial_x H \mathrm{d}t + Q \mathrm{d}W_b(t) \tag{3.114}$$

with the initial condition

$$p(T) = -\partial_x \Phi(x(T)). \tag{3.115}$$

This we can enter into the position SDE (3.109) resulting in

$$\mathrm{d}x(t) = \frac{p(x(t),t)}{m}\mathrm{d}t + \sqrt{\frac{\hbar}{m}}\mathrm{d}W_q(t). \tag{3.116}$$

Now, because we describe conservative Brownian motion, we have time invariance, meaning we can split the forward and backward equations into equivalent forward and backward equations, i.e. for the Hamiltonian

$$H_b(t,x,v_q,p,Q) = -\frac{m}{2}v_q^2 + V(x) + p_b v_q - \sqrt{\frac{\hbar}{m}}\frac{1+i}{2}Q_b \tag{3.117}$$

$$H_f(t,x,v_q,p,Q) = -\frac{m}{2}v_q^2 + V(x) + p_f v_q - \sqrt{\frac{\hbar}{m}}\frac{1+i}{2}Q_f. \tag{3.118}$$

Due to (3.112), we get

$$p_f = p_b = p = mv_q \tag{3.119}$$

we get also two equivalent equations for the momentum,

$$\mathrm{d}p_b(t) = -\partial_x V(x)\mathrm{d}t + Q_b\mathrm{d}W_b(t), \tag{3.120}$$

$$\mathrm{d}p_f(t) = -\partial_x V(x)\mathrm{d}t + Q_f\mathrm{d}W_f(t), \tag{3.121}$$

which can be added together:

$$\mathrm{d}p(t) = -\partial_x V(x)\mathrm{d}t + \frac{1}{2}(Q_f\mathrm{d}W_f(t) + Q_b\mathrm{d}W_b(t)). \tag{3.122}$$

We still need to know, what $Q_{f/b}$ is. Looking at $p(x(t),t)$ as a function of a stochastic process $x(t)$ we can use the Itô formula which yields another SDE that describes the process $p(t)$. We only have to remind ourselves that we are in complex space with

$$\mathrm{d}W_q(t)\mathrm{d}W_q(t) = -i\mathrm{d}t, \tag{3.123}$$

which we have to use as a factor for the third term of the drift part of the Itô formula, resulting in our SDE

$$\mathrm{d}p(x(t),t) = \left[\partial_t + v_q\partial_x - i\frac{\hbar}{2m}\partial_x^2\right]p(x(t),t)\mathrm{d}t + \sqrt{\frac{\hbar}{m}}\partial_x p(x(t),t)\mathrm{d}W_q(t). \tag{3.124}$$

Comparing the diffusion terms from the two SDEs yields

$$Q_b = \sqrt{\hbar m}(1-i)\partial_x v_q(x,t) \tag{3.125}$$

$$Q_f = \sqrt{\hbar m}(1+i)\partial_x v_q(x,t) \tag{3.126}$$

for the costate processes and also

$$-\partial V_x = \left[\partial_t + v_q\partial_x - i\frac{\hbar}{2m}\partial_x^2\right]m(v(x,t) - iu(x,t)). \tag{3.127}$$

Thus we have a complex SDE for the momentum,

$$\mathrm{d}p(x,t) = -\partial_x V(x)\mathrm{d}t + \sqrt{\frac{\hbar}{m}}\partial_x p(x,t)\mathrm{d}W_q(t). \tag{3.128}$$

which was also derived by [Pav95]. Splitting (3.127) into a real and imaginary part,

$$-\frac{1}{m}\partial V_x = \partial_t v + v\partial_x u - \frac{\hbar}{2m}\partial_x^2 iu(x,t)), \tag{3.129}$$

$$0 = \partial_t u + u\partial_x v + \frac{\hbar}{2m}\partial_x^2 v(x,t), \tag{3.130}$$

where the first equation is equal to (3.66) and both are equivalent to the Madelung equations

$$\partial_t \rho(x,t) + \frac{1}{m}\partial_x(\rho(x,t)\partial_x S(x,t)) = 0, \tag{3.131}$$

$$v(x,t) = \frac{1}{m}\partial_x S(x,t), \tag{3.132}$$

$$\partial_t S(x,t) + \frac{1}{2m}(\partial_x S(x,t))^2 + V(x,t) - \frac{\hbar^2}{2m}\frac{\partial_x^2\sqrt{\rho(x,t)}}{\sqrt{\rho(x,t)}} = 0. \tag{3.133}$$

To get the formulation in real space, we have a forward equation for the position and a backward equation in momentum

$$\mathrm{d}x(t) = (v+u)\mathrm{d}t + \sqrt{\frac{\hbar}{m}}\mathrm{d}W_f(t), \tag{3.134}$$

$$\mathrm{d}p(t) = m\mathrm{d}(u - iu) = -\partial_x V(x)\mathrm{d}t + \sqrt{\hbar m}(1+i)\partial_x(v - iu)\mathrm{d}W_b(t), \tag{3.135}$$

where we again split the momentum into the real and imaginary part

$$\mathrm{d}v = -\frac{1}{m}\partial_x V(x)\mathrm{d}t + \sqrt{\frac{\hbar}{m}}\partial_x(u+v)\mathrm{d}W_b(t), \tag{3.136}$$

$$\mathrm{d}u = \sqrt{\frac{\hbar}{m}}\partial_x(u-v)\mathrm{d}W_b(t) \tag{3.137}$$

$$\tag{3.138}$$

and adding them together

$$\mathrm{d}\tilde{p}(t) = \mathrm{d}(v+u) = -\partial_x V(x)\mathrm{d}t + 2\sqrt{\hbar m}\partial_x u\mathrm{d}W_b(t) \tag{3.139}$$

to get the real-valued SDE for the momentum process. Here we can see, that classical behavior is recovered in the classical limit

$$\frac{\hbar}{m} \to 0 \Rightarrow u \to 0 \tag{3.140}$$

with

$$\mathrm{d}x(t) = v(t)\mathrm{d}t = \partial_p H\mathrm{d}t, \tag{3.141}$$

$$\mathrm{d}p(t) = -\partial_x H\mathrm{d}t \tag{3.142}$$

being the classical Hamilton equations(3.17, 3.18). Taking the expectation values, we recover the Ehrenfest theorems

$$\frac{\mathrm{d}\langle x(t)\rangle}{\mathrm{d}t} = \frac{\langle p(t)\rangle}{m}, \tag{3.143}$$

$$\frac{\mathrm{d}\langle p(t)\rangle}{\mathrm{d}t} = \langle F(x(t))\rangle. \tag{3.144}$$

Equations (3.134) and (3.139) are called the non-stationary quantum Hamilton equations due to their similarities to the classical Hamilton equations.
For the stationary ground state we have to perform a different derivation, because the math in [BG10] is only valid for non-stationary problems. We follow [ØS14] and [Øks10]. In the stationary case, the action function separates into

$$S(x,t) = S_0(x) - Et \tag{3.145}$$

with the position-dependent part disappearing for ground states. Thus, the drift velocity is

$$v = 0 \tag{3.146}$$

for $S_0(x) = 0$. Because of that, we only have one cost functional

$$J[u] = \max_u \mathbf{E}\bigg[\int_0^T \Big(-\frac{m}{2}u^2 - V(x(t))\Big)\mathrm{d}t + S(x_0)\bigg] \tag{3.147}$$

giving the controlled forward and backward SDEs

$$\mathrm{d}X = u(x(t))\mathrm{d}t + \sqrt{\frac{\hbar}{m}}\mathrm{d}W_f, \tag{3.148}$$

$$\mathrm{d}X = -u(x(t))\mathrm{d}t + \sqrt{\frac{\hbar}{m}}\mathrm{d}W_b. \tag{3.149}$$

From the Pontryagin maximum principle (see Appendix C.1), which introduces an additional control u, we obtain the Hamiltonian

$$H(x,u,\lambda,p,q) = -\frac{m}{2}u^2 - V(x) + \lambda u + pu + \sqrt{\frac{\hbar}{m}}q \tag{3.150}$$

with λ being a forward adjunct process to the backward drift, p being a backward adjunct process to the forward drift and q being the adjunction process to the backward Wiener process. The λ process follows the equation

$$\mathrm{d}\lambda = 0 \tag{3.151}$$

and with the initial condition $\lambda(0) = 0$ it is zero for all times. The p process is governed by the backward equation

$$\mathrm{d}p(t)) = \frac{\mathrm{d}V(x)}{\mathrm{d}x} + q(t)\mathrm{d}W_b(t) \tag{3.152}$$

with $\mathrm{d}t < 0$. We can use the maximum principle again

$$\partial_u H = 0, \tag{3.153}$$

which leads to

$$mu = p + \lambda = p. \tag{3.154}$$

To obtain an equation for the q process, we again employ the Itô formula to get

$$0 = \partial_t u + u\partial_x u + \frac{\hbar}{2m}\partial_x^2 u - \frac{1}{m}\partial_x V, \tag{3.155}$$

$$q(t) = \sqrt{\hbar m}\partial_x u(x), \tag{3.156}$$

with the first one being the equivalent of the time-dependent Schrödinger equation, which leads to a set of coupled forward/backward stochastic differential equations (FBSDEs)

$$\mathrm{d}x(t) = u(x(t))\mathrm{d}t + \sqrt{\frac{\hbar}{m}}\mathrm{d}W_f, \tag{3.157}$$

$$\mathrm{d}u(t) = \frac{1}{m}\frac{\mathrm{d}}{\mathrm{d}x}V\mathrm{d}t + \sqrt{\frac{\hbar}{m}}\partial_x u(x(t))\mathrm{d}W_b(t), \tag{3.158}$$

which are the quantum Hamilton equations (QHE) for the ground state.

The non-stationary quantum Hamilton equations do not simplify to the stationary ones by simply setting the current velocity to zero. One still has a factor of 2 in the diffusion term, that vanishes in the stationary case. This discrepancy should be a concern for future research.

3.5 Solving the Quantum Hamilton Equations with the Bender/Steiner Method

The quantum Hamilton equations can be solved numerically without using the Schrödinger equation with the iterative scheme proposed in [BS12] and first used in [KGP16] for the quantum Hamilton equations. One starts with an initial guess of the osmotic velocity $u^0(x)$. The convergence of the algorithm doesn't depend on the initial guess, so one can choose $u^0(x) = 0$. This guess is then used to solve the forward SDE numerically. This solution is then used to find an update of the osmotic velocity. This solution is then used to solve the forward SDE etc.. This iterative procedure is repeated until convergence is reached. The forward solutions and backward solutions are generated multiple times per iteration to achieve convergence faster. The general scheme is as follows:

- initialize $u^0(x) = 0$
- repeat N times or until convergence
  - repeat over all time steps and over all sample paths
    - $x^m(0) = x_0$
    - propagate m-th particle forward in time: $x_1^m = x_0^m + u(x_0)\Delta t + \sigma\Delta W$
    - update $u(x)$
  - average over all u-updates

The forward propagation is straightforward by using, for example, the Euler-Maruyama or the Heun scheme. However, there are different methods for updating the osmotic velocity.

One is solving the ordinary differential equation (3.155) iteratively,

$$0 = u^{j-1}\frac{\mathrm{d}u^j(x)}{\mathrm{d}x} + \frac{1}{2}\frac{\mathrm{d}^2u^j(x)}{\mathrm{d}x^2} - \frac{\mathrm{d}V(x)}{\mathrm{d}x}, \tag{3.159}$$

with the boundary conditions

$$u^j(X_T) = u^{j-1}(X_T), \tag{3.160}$$

$$\left.\frac{\mathrm{d}u^j(x)}{\mathrm{d}x}\right|_{x=X_T} = \left.\frac{\mathrm{d}u^{j-1}(x)}{\mathrm{d}x}\right|_{x=X_T}. \tag{3.161}$$

There, j is the iteration step for each forward realization. Eventually, one averages over each solution. This method is only valid for the stationary case.
Another method is to directly evaluate the backward SDE. For this, we take each forward propagation X_t of the position, start at $t = T$ and propagate the osmotic velocity path backward in time:

$$u^j_{t_i} = u^j_{t_{i+1}} - \left.\frac{\mathrm{d}V(x)}{\mathrm{d}x}\right|_{x=X_{t_i}}\Delta t_i - \left.\frac{\mathrm{d}u^j(x)}{\mathrm{d}x}\right|_{x=X_{t_i}}\Delta W_i, \tag{3.162}$$

with the initial conditions

$$u^j_T = u^{j-1}(X_T), \tag{3.163}$$

$$\left.\frac{\mathrm{d}u^j(x)}{\mathrm{d}x}\right|_{x=X_T} = 0 \tag{3.164}$$

(Δt_i and ΔW_i discretizations of the time and the Wiener process respectively). Sometimes taking the gradient of the osmotic velocity does not work. One then has to replace the gradient of u in the backward equation by the adjoint process $q(x)$. This process is based on conditional expectation (see [Kö18] for more details). One then obtains the iterative backward propagation

$$u^j_{t_N} = u^{j-1}(X^j_{t_N}), \quad q^j_{t_N} = 0, \tag{3.165}$$

$$q^j_{t_i} = \frac{1}{\Delta t_i}\alpha^j_i\eta(X_{t_i}), \quad \alpha^j_i = \frac{1}{M}\sum_{j=1}^{M}\eta(X_{t_i})u_{t_{i+1}}\Delta W_i, \tag{3.166}$$

$$u^j_{t_i} = u^j_{t_{i+1}} - \left.\frac{\mathrm{d}V(x)}{\mathrm{d}x}\right|_{x=X_{t_i}}\Delta t_i - q^j_{t_i}\Delta W_i, \tag{3.167}$$

with α being the functional basis of the conditional expectation [GLW05] and η being a discretion vector of the space to make the equations suitable for computation and M being the total number of sample paths. The update after each iteration is done by calculating the average over all u^j-realizations.
Another option is replacing q by using another approach concerning conditional expectation [Bey18]. The forward process implies the conditional expectations

$$\mathbb{E}[f(t_i)\Delta W(t_i)|x(t_i)] = 0, \tag{3.168}$$

$$\mathbb{E}[g(x(t_i))|x(t_i)] = g(t_i) \tag{3.169}$$

for any functions $f(t), g(x(t)) \in \mathbb{R}^n$ which are adapted to the forward path. Taking the conditional expectation on (3.167) one obtains

$$u^{\pi,j}(x^{\pi,j}(t_i)) = -\partial_x V(x^{\pi,j}(t_i)\Delta t + \mathbb{E}[u^{\pi,j}(x^{\pi,j}(t_{i+1}))|x^{\pi,j}(t_i)] \tag{3.170}$$

with π indicating sample path and j indicating the time step. Consider that $\mathbb{E}[g(x(s))|x(t)] \neq g(t)$ if $s > t$. (3.170) can be approximated by using N sample paths

$$u^{\pi,j,l}(t_i) = \frac{1}{\#x^l(t_{i+1})} \sum_{n=1}^{N} [\delta(x^l(t_i) - x^n(t_{i+1}))u^n(t_{i+1}] - \partial_x V(x^{\pi,l,j}(t_i)\Delta t, \tag{3.171}$$

where l represents the sample path out of N which is analyzed and $\#x^l(t_{i+1})$ is the number of times the position $x^l(t_{i+1})$ was visited by the other sample paths.

As one might guess, the implementation of appraoches is challenging, prone to errors and needs a lot of computational resources. It suffers also heavily from the curse of dimensionality. The parameters $N, M, \eta, \Delta t$ also have to be chosen very carefully to guarantee convergence. A discussion of the parameters can be found in [Kö18].

3.5.1 Example: One-Dimensional Harmonic Oscillator

To solve the harmonic oscillator, we use the gradient approach described in the previous section, resulting in the following algorithm

$$x_1 \leftarrow x_0 + u^{(old)}(x_0)\Delta t + \sigma \Delta W, \tag{3.172}$$

$$u^{(new)}(x_0) \leftarrow u^{(old)}(x_0) - \nabla V(x_1)\Delta r - \sigma[\partial_x u^{(old)}(x_0)(x_1)]\Delta W. \tag{3.173}$$

The resulting osmotic velocity, shown in Fig. 3.3. One can see, that we get a very close to correct result for $x \in (-2, 2)$. The areas a little bit further out have started to converge but do so much slower and the areas from $|x| > 4$ have not started to converge at all, being still at the the initial condition. From the osmotic velocity one can calculate the probability density using a formula described and derived in [Kö18]:

$$\rho(x) = c \exp\left\{2 \int_{-\infty}^{x} u(x')\mathrm{d}x\right\}, \tag{3.174}$$

$$c = \Big[\int_{-\infty}^{\infty} \exp\left\{2 \int_{-\infty}^{x} u(x')\mathrm{d}x\right\}\mathrm{d}x\Big]^{-1}. \tag{3.175}$$

The ground state energy then reads

$$E_0 = \int_{-\infty}^{\infty} \left(\frac{1}{2}u(x)^2 + V(x)\right)\rho(x)\mathrm{d}x \tag{3.176}$$

which gives the numerical value $E_0 = 0.5611$ (correct value: 0.5). Much better results can be obtained by fine-tuning the parameters of the algorithm [Kö18].

3.6 Exited States

In section 3.3.1 we have seen that the Nelson formulation is equivalent to the Madelung equations. But the Madelung equations are only equivalent to the Schrödinger equation

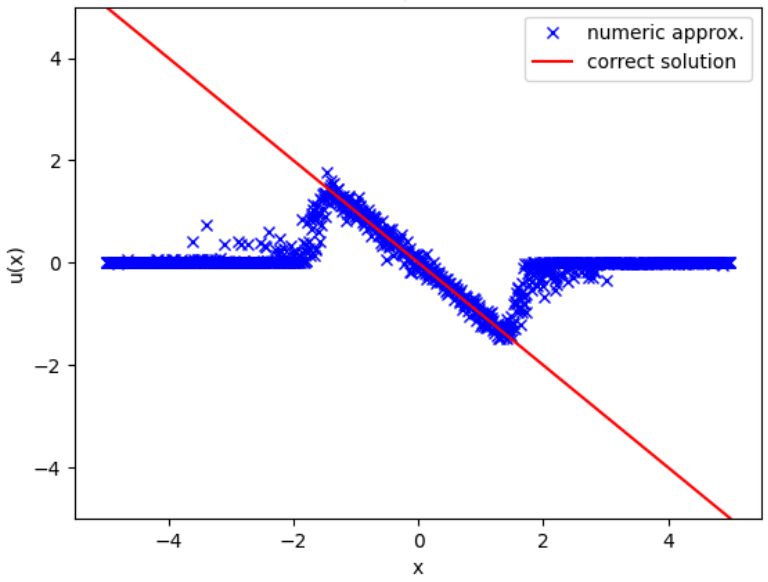


FIGURE 3.3: Osmotic velocity and corresponding probability density for the 1-D harmonic oscillator using $k = \sigma = m = 1$ and atomic units a discrete time step of $\Delta t = 0.5$, 10000 time steps, 10 sample paths M and $N = 10$ episodes. The probability density was calculated using (3.174) and the trapezoidal method. Numerical ground state energy was calculated using (3.176) $E_0 = 0.5611$ (correct value: 0.5).

for node-less states, e.g. the ground states of quantum systems. The standard method of calculating exited states for the stationary case in Nelson mechanics is given in [Pat18]. Using atomic units here for simplicity, the ground state Hamiltonian

$$H_0 = -\frac{1}{2}\partial_x^2 + V_0(x) \tag{3.177}$$

for bound states, which obey the Schrödinger equation

$$H_0\Psi_0^n = E_0^n\Psi_0^n \tag{3.178}$$

can be factorized with the two adjoint operators

$$Q^- = \frac{1}{\sqrt{2}}(\partial_x + W(x)), \tag{3.179}$$

$$Q^+ = \frac{1}{\sqrt{2}}(-\partial_x + W(x)), \tag{3.180}$$

resulting in

$$H_0 = Q^+Q^- + \varepsilon \tag{3.181}$$

with an yet undetermined constant ε. Applied to the ground state,

$$(Q^+Q^- + \varepsilon)\Psi_0^0 = \left(-\frac{1}{2}\partial_x^2 + \frac{1}{2}W^2 - \frac{1}{2}\partial_x W + \varepsilon\right)\Psi_0^0 = H_0\Psi_0^0, \tag{3.182}$$

and compared to the Schrödinger equation, results in the following Riccati-equation (see Appendix D)

$$W^2 - \partial_x W = 2(V_0 - \varepsilon). \tag{3.183}$$

Substituting $W(x) = -\frac{\partial_x \phi}{\phi}$ leads to

$$-\frac{1}{2}\partial_x^2 \psi + (V_0 - \varepsilon)\psi = 0 \tag{3.184}$$

which is satisfied if

$$W(x) = \frac{\partial_x \Psi_0^0}{\Psi_0^0} = -\partial_x \ln \Psi_0^0, \tag{3.185}$$

$$\varepsilon = E_0^0. \tag{3.186}$$

The factorized Hamiltonian has a super-symmetric partner with reversed order of the adjoint operators

$$H_1 = Q^- Q^+ + E_0^0, \tag{3.187}$$

which results in

$$H_1 = -\frac{1}{2}\partial_x^2 + V_1(x) \tag{3.188}$$

$$V_1(x) = V_0(x) - \partial \psi_0^0. \tag{3.189}$$

Thus the adjoint operators read

$$Q^+ = \frac{1}{\sqrt{2}}(-\partial_x - \partial_x(\ln \Psi_0^0)), \tag{3.190}$$

$$Q^- = \frac{1}{\sqrt{2}}(\partial_x - \partial_x(\ln \Psi_0^0)). \tag{3.191}$$

One can take the new Hamiltonian H_1 and calculate the ground state wave function, which solves the corresponding Schrödinger equation for Ψ_1^0. From this one can obtain the corresponding adjoint operators by replacing the ground state Ψ_0^0 by Ψ_1^0, resulting in

$$Q_1^+ = \frac{1}{\sqrt{2}}(-\partial_x - \partial_x(\Psi_1^0)), \tag{3.192}$$

$$Q_1^- = \frac{1}{\sqrt{2}}(\partial_x - \partial_x(\Psi_1^0)). \tag{3.193}$$

From this one can calculate H_2 with the corresponding wave functions and energies and so on. Thus it is possible to calculate every exited state from the ground state for all systems, which have a bounded ground state (see Fig. 3.4).
But how do we obtain the exited states in the Nelson formulation? Starting with the ground state Hamiltonian

$$H_0 = -\frac{1}{2}\Delta + V_0(x) \tag{3.194}$$

results in the stationary case in the QHE

$$\mathrm{d}X = u\mathrm{d}t + \sqrt{\frac{\hbar}{m}}\mathrm{d}W_f, \qquad X(t=0) = x_0 \in \mathbb{R}^{3N}, \tag{3.195}$$

$$\mathrm{d}u(t) = \frac{1}{m}\nabla V \mathrm{d}t + \frac{q(t)}{m}\mathrm{d}W_b(t), \qquad u(X(T)) = 0. \tag{3.196}$$

From this, we can calculate the ground state solution u_0, and thus ψ_0^0 and the partner Hamiltonian

$$H_1 = -\frac{1}{2}\Delta + V_1(x). \tag{3.197}$$

Using eq.(3.189) with the connections to the Schrödinger formalism (3.91) and (3.93) we can get $V_1(x)$ by

$$\psi_0^0 = e^{-\tilde{R}_0(x)}, \tag{3.198}$$

$$\partial_x \tilde{R}_0(x) = -\partial_x \ln \psi_0^0 = -u_0, \tag{3.199}$$

$$V_1 = V_0 - \partial_x u, \tag{3.200}$$

which results in the new QHE for the first exited state

$$\mathrm{d}X = u_1\mathrm{d}t + \sqrt{\frac{\hbar}{m}}\mathrm{d}W_f, \qquad X(t=0) = x_0 \in \mathbb{R}^{3N}, \tag{3.201}$$

$$\mathrm{d}u_1(t) = \frac{1}{m}\nabla V_1 \mathrm{d}t + \frac{q(t)}{m}\mathrm{d}W_b(t), \qquad u(X(T)) = 0. \tag{3.202}$$

Like in the Schrödinger formalism, one calculates all bound exited states from the ground state using the iterative scheme

$$\begin{aligned} & H_0 \xrightarrow{QHE} u_0 \\ \text{calculate} \quad & H_1 = H_0 - \frac{\mathrm{d}}{\mathrm{d}x}u_0 \\ & H_1 \xrightarrow{QHE} u_1 \\ & .... \end{aligned}$$

3.7 Tunneling and Tunneling Times

A huge advantage of stochastic mechanics is that time becomes a measurable entity. In wave mechanics, there is no time measurement operator, thus time is not an observable there. Additionally, because we do not view quantum systems as closed, rather as open systems, there is no need for the concept of tunneling to describe barrier crossings. Since energy fluctuates, the particle does not tunnel through a barrier, but goes over the barrier. This process is driven by quantum fluctuations.

Previous attempts in wave mechanics to determine the duration of tunneling processes did predict anything from instantaneous tunneling to superluminal speeds to traveling of particles with the speed of light. Stochastic mechanics instead gives simple explanations and tools to obtain results that fit with reality.

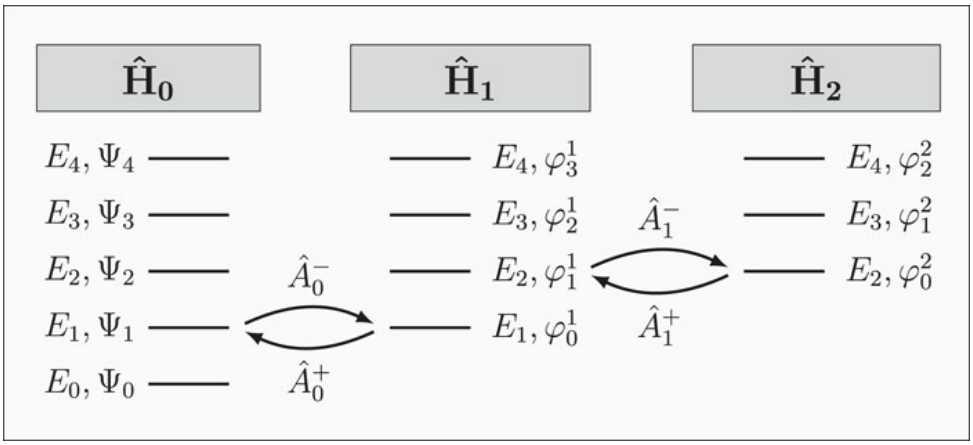


FIGURE 3.4: A Hamiltonian and its partner Hamiltonians for the the first and second exited states are related to each other by way of the ladder operators. The ground state energies for the first and second partner Hamiltonian are equal to the corresponding exited energies of the original Hamiltonian (figure from [Kö+18]).

The first method is simulating many sample paths X_t and measure for each path, when a barrier is crossed. Then one calculates the mean first passage time t_{mfpt} by taking the arithmetic average over all first passage times.
Another method is using Kramer's theory, sometimes called Kramer's problem [PB13]. Kramer's problem can be formulated in one sentence: "If we have a particle undergoing Brownian motion in a metastable (local minimum) or equilibrium state (global minimum) in a potential, how long does it on average take to move a particle from one of these into another local or global minimum of the potential?" H.A. Kramer developed this method to describe chemical reactions [Kra40]. Using the Fokker-Planck equation one can derive expressions for the mean first passage time

$$t_{mfpt}(x) = \frac{2m}{\hbar} \int_x^{x_t} \frac{\mathrm{d}x'}{\rho_0(x')} \int_{-\infty}^{x'} \rho_0(x'')\mathrm{d}x'' \tag{3.203}$$

for potentials, where there are areas, in which the particle does not interact with the potential and

$$t_{mfpt}(x) = \frac{2m}{\hbar} \int_{-\infty}^{-x_t} \tilde{\rho}_0(x) \int_x^{x_t} \frac{\mathrm{d}x'}{\rho}_0(x') \int_{-\infty}^{x'} \rho_0(x'')\mathrm{d}x'' \tag{3.204}$$

where $\tilde{\rho}_0(x)$ is the renormalized probability density from $\rho_0(x)$ averaged over all starting points between $-\infty$ and x'.

3.8 Stochastic Mechanics and Wigner's Friend

In [Wig95], Eugene Wigner proposed a thought experiment to show the paradoxical nature of conventional quantum mechanics, which results from the contradictory nature of the wave function collapse and the time evolution dictated by the Schrödinger equation.
Consider a qubit system, a two-state quantum system S with both states, labeled $|0\rangle$ and $|1\rangle$,

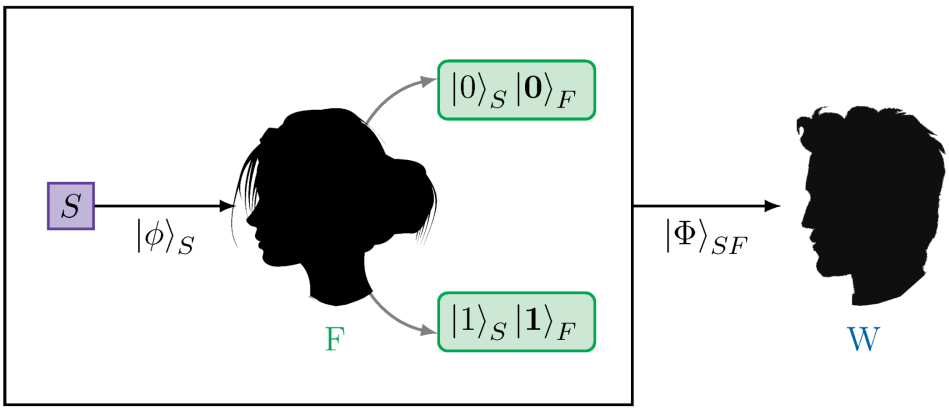


FIGURE 3.5: Sketch of the Wigner's friend thought experiment(taken from [Bau23]). While the friend observes a collapse of the wave function, from Wigner's point of view, his friend and the system entangle with one another in accordance with the Schrödinger equation.

being possibly measured with equal probability, ergo

$$|\Psi\rangle_S = \frac{1}{\sqrt{2}}(|0\rangle_S + |1\rangle_S). \tag{3.205}$$

This system is observed by two observers colloquially labeled Wigner W and his Friend F. F is inside the laboratory, while W waits outside. Now, if F measures the system, the system and friend can be described by the entangled wave function

$$|\Psi\rangle_{SF} = \frac{1}{\sqrt{2}}(|0\rangle_S \otimes |0\rangle_F + |1\rangle_S \otimes |1\rangle_F), \tag{3.206}$$

where $|0/1\rangle_F$ denotes that F observes $|0/1\rangle_S$ [LB21]. If Wigner now enters the laboratory and asks his friend, what was measured, the friend will answer $|0\rangle$ or $|1\rangle$. In conventional quantum mechanics, the wave function of the lab collapsed. However, for the friend the systems wave function alerady collapsed with their measurement, the entanglement did not occur. This is a paradox, because it is not clear, when the wave function of the system collapsed.
Other interpretations of quantum mechanics resolve this paradox in some ways. The many-worlds interpretation does not have wave function collapse at all but the wave functions includes more and more entangled systems. In the present case, Wigner would entangle with his friend and the system, resulting in

$$|\Psi\rangle_{SF} = \frac{1}{\sqrt{2}}(|0\rangle_S \otimes |0\rangle_F \otimes |0\rangle_W + |1\rangle_S \otimes |1\rangle_F \otimes |1\rangle_W). \tag{3.207}$$

In Bohmian mechanics, the paradox is resolved by having the system following classical trajectories, that are however unknown due to the uncertainty in the initial conditions. The result of the measurement by the friend was set by those initial conditions. Wave function collapse is thus not needed here either [LB21].
Stochastic mechanics resolves this paradox in a similar fashion. Which stochastic trajectory the particle takes before the measurement is unknown. What the friend measures does not result in an entangled state of them and the system. Firstly, because Nelson's stochastic mechanics does not need a wave function. But secondly, because stochastic mechanics has a clear and continuous separation of classical mechanics and quantum mechanics due to the

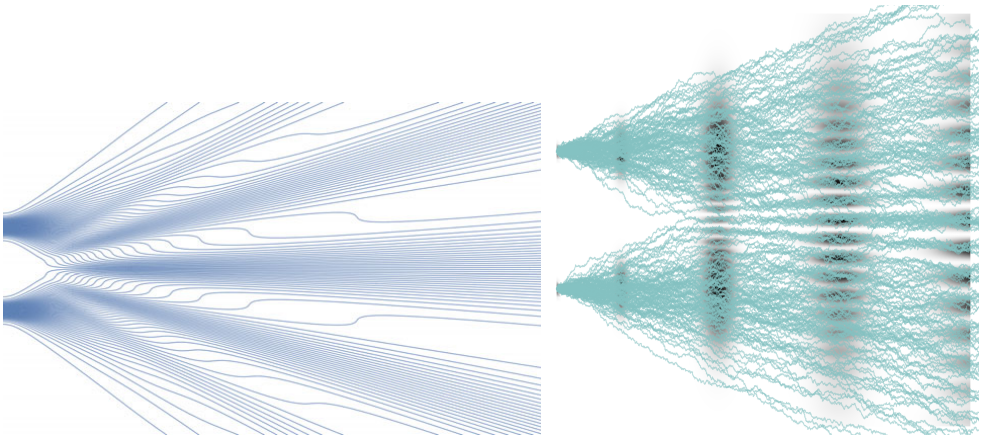


FIGURE 3.6: Comparison between Bohmian trajectories [Nor17] and stochastic trajectories from [Bey23] for the double-slit experiment. Both theories reproduce the results of quantum mechanics.

mass dependence of the diffusion term

$$\sigma \propto \frac{1}{\sqrt{m}}. \tag{3.208}$$

The quantumness of the friend and Wigner are thus trivially inconsequential and negligible and the paradox is resolved.

3.9 Equivalency to other Formulations of Quantum Mechanics

We have already shown that stochastic mechanics is equivalent to the Schrödiner equation and the Madelung equations.
In Bohmian mechanics, the particle "surfes" on a wave that obeys the Schrödinger equation. Its trajectory is related to the phase of the wave function with the differential equation

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\hbar}{m}\frac{\partial S(x,t)}{\partial x}\bigg|_{x=X(t)}, \tag{3.209}$$

which can also be written as an equation in term of the wave function as

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\hbar}{m}\,\mathrm{Im}\left(\frac{1}{\Psi(x,t)}\frac{\partial\Psi(x,t)}{\partial x}\right)\bigg|_{x=X(t)} = \mathrm{Re}\left(\frac{\hat{p}\Psi(x,t)}{\Psi(x,t)}\right)\bigg|_{x=X(t)}, \tag{3.210}$$

resulting in smooth deterministic particle trajectories [Nor17]. A comparison between the different kinds of trajectories can be seen in Fig. 3.6. From equation (3.92), one can see that the velocity of the Bohmian picture is the same as the current velocity of stochastic mechanics. The Bohmian trajectories follow the quantum potential

$$V_Q = -\frac{\hbar}{2m}\frac{\Delta|\Psi|}{|\Psi|}, \tag{3.211}$$

which can be expressed in terms of the osmotic velocity [Bey23]

$$V_Q = -\frac{m}{2}u^2 - \frac{\hbar}{2}\nabla u \tag{3.212}$$

with expectation value

$$\mathbb{E}[V_Q] = \mathbb{E}\left[\frac{m}{2}u^2\right]. \tag{3.213}$$

The difference between stochastic mechanics and Bohmian mechanics is that the wave function is not needed in the former. The osmotic velocity and the current velocity are the source of quantum effects in stochastic mechanics, not the wave function.

The description of Nelson mechanics is equivalent to Feynman's path integral formulation of quantum mechanics [FHS10]. The paths that are integrated over in Feynman's path integral formulation of quantum mechanics

$$\Psi(x,t) = \frac{1}{C}\int e^{\frac{i}{\hbar}\int_t^{t'}\mathcal{L}(x,\dot{x},t')\mathrm{d}t''}\Psi(x,t')\mathcal{D}x \tag{3.214}$$

(C the normalization factor, $\mathcal{L}$ the Lagrangian of the problem, $\mathcal{D}x$ denotes the integral over all possible paths) are the stochastic paths in this description. This is proven in [Pav00].

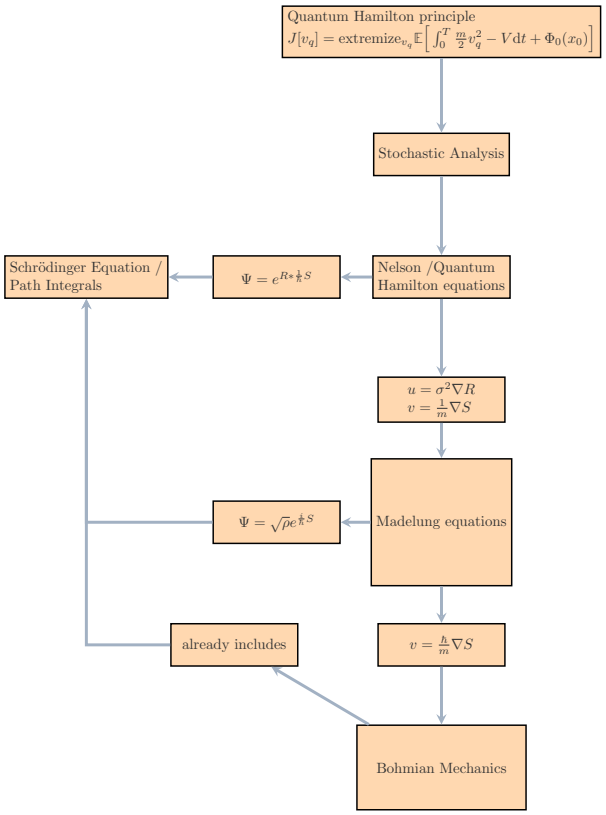


FIGURE 3.7: Relations of different formulations of quantum mechanics.

3.10 Connection to other Interpretations of Quantum Mechanics

The many-worlds interpretation regards the wave function of the Schrödinger equation as something existing (ontological). In contrast to the Copenhagen interpretation, where the wave function collapses if a measurement with a classical object occurs, in the many-worlds interpretation the interaction with the measurement apparatus is described by a single wave function which includes the quantum object and the apparently "classical" measurement apparatus. This can then be extended outwards until the entire universe is described by a single wave function, the universal wave function. The fact that we as human beings are living in a classical world is due to the fact that we can only observe a single superposition within the universal wave function. E.g. if we measure the spin of an electron, that was previously unknown, our measurement apparatus can either say the electron is spin up or spin down. There is a superposition of the wave functions of purely spin *up* and purely spin *down*. If

we measure the spin and get a measurement of *up*, the superposing wave function extends to us an we are now part of the spin *up* superposition. But the spin *down* superposition did not cease to exist, like in the Copenhagen interpretation, but it still exists in another, for us unobservable, part of the universal wave function. Of course there is another world, another part of the wave function, where we measured spin *down*, but this world or part of the universal wave function is unobservable to us.
On the surface, this interpretation seems completely different from Nelson's random paths. But let us consider a random path of a free electron in a cloud chamber. In Nelson mechanics, the electron obeys the equations of motion of a free particle of classical mechanics disturbed by the noise of the Wiener process. What path the electron takes if it was to go through the cloud chamber is random. In the many-worlds interpretation, the universe splits into different superpositions, one for each possible path, and we observe only one path because we become part of the different superpositions ourselves. This way one could interpret the Wiener process $\sigma\Delta W(t_i)$ as a discrete universe splitter. Nelson mechanics is usually seen as a hidden variables theory (and all the hidden variables amount to the Wiener process in the QHEs) and the many worlds interpretation is usually seen as not being a hidden variables theory. This means they are usually seen as epistemologically opposed. The above suggests that the many-worlds theory can be formulated without the use of wave functions.

Stochastic Mechanics is as a remark also connected with interpretations of quantum mechanics that require wave function collapse like the Copenhagen interpretation. This can be shown by the use of Bayes' theorem [Pav99].

3.11 Controversies

The stochastic mechanics description of quantum mechanics is not very popular. The reasons for this are manifold. One is the grip that the Copenhagen interpretation still has on the way quantum mechanics is taught in universities. Another one is a presuming misunderstanding of Bell inequalities in the physics community. It is commonly believed that their violation disproves all hidden variable theories, even though they only disprove local, or Bell-separable, hidden variable theories, a mistake even made by the Nobel committee in 2022 [Sci22]). A historical reason is John von Neumanns "No-hidden-variables-proof" which he published in his foundational book on quantum mechanics [Von32], even though it was proven false in 1935 by Grete Herrmann [Her35; See16], her critique was ignored by almost all physicists at the time.
However, there are still valid points of critique in regard to the Nelson picture. Wallström showed that for systems with non-zero angular momentum, one needs additional ad-hoc conditions to get the same exited states as the Schrödinger equation [Wal94]. He derived this result also for the Madelung equations.
Another point of criticism is the origin of the background field. Many different ones have been proposed. Nelson proposed cosmic background radiation as a possible candidate [Nel85], and gravitational waves were proposed in [Red17]. The specific nature will not be discussed further in this thesis.

Chapter 4

Genetic Algorithms

Genetic algorithms are a class of algorithms under the umbrella of evolutionary algorithms. The general idea is to mimic natural evolution in a simplified way for solving optimization problems. The description of the basic concepts of evolutionary algorithms will be taken from [Spe00; Gol89; Mit98].
In what follows, we take concepts and principles from biology and transform them to our needs. To underline this origin, we will use biological terminology.
The main principle of genetic algorithms is "survival of the fittest". Fitness is an abstract concept in nature and is still, after almost 200 years after Herbert Spencer coined the term (Charles Darwin popularized the term in the fifth edition of his foundational text "The Origin of Species"), not entirely understood and is still subject of research in evolutionary biology. Here we will introduce a first simplification. To find the fittest individuals, we need to compare them, ergo we need a scalar function f, called fitness function. This function measures the fitness of each individual in a population. The population is the set of all individuals.

4.1 Genotype, Phenotype

For this we need to first describe our individuals, genotype and phenotype. In nature, genotype is the sequence of the four nucleotide bases, adenine, guanine, cytosine, thymine, in the entirety of the genome of an organism. This parametrization is simplified, ergo we do not use base-4 but base-2, 0's and 1's. This parametrization leads to a phenotype, the measurable characteristics of the individual. This phenotype is what is used to calculate the fitness of each individual and is what is of interest in genetic algorithms. To repopulate after each selection, there are two main methods, Recombination and Mutation.

4.2 Selection and Survival

The other part of the Darwinian principle is survival, meaning that only a set number of individuals survives each generation. Each generation, we calculate f of each individual and then pick survivors. Fitness and probability of survival are closely linked to one another: the higher the fitness, the higher the probability of survival. This could be simply achieved by picking the fittest ones and discarding the rest. But one could also introduce a probability measure $p(f)$. Again, there are a lot of choices to be made. A simple method would be for example to divide the fitness of the $i-$th individual by the sum of fitness of the entire

population of size N,

$$p_i = \frac{f_i}{\sum_{j=1}^{N} f_j}, \tag{4.1}$$

which works if the possible fitness values are all positive semi-definite. Then randomly pick survivors according to the possibilities.
One could also implement a rank choice system in which the individuals are ranked with regard to their fitness. Then we choose a probability constant p_c, for which the fittest individual survives. If the fittest one is not chosen to survive we go down the list until the last one ergo

$$\begin{aligned} p_1 &= p_c \\ p_2 &= (1-p_c)p_c \\ &\vdots \\ p_i &= (1-p_c)^{i-1}p_c \\ &\vdots \\ p_N &= (1-p_c)^{N-1}. \end{aligned} \tag{4.2}$$

There are many other methods of picking survivors.It should be kept in mind that the higher the fitness, the higher the probability of survival.

4.3 Repopulation

After the survivors are picked, the population is repopulated.

4.3.1 Recombination/Crossover

Recombination, also called crossover, is the simplified version of biological, sexual reproduction. We take the chromosomes of two surviving individuals and recombine them in some way. In theory, we could take chromosomes from more than one individual, however this is rarely done. This recombination can be done by, for example, splitting the genotype of the parents into parts at the same locations and then mixing them up to get two individuals with mixed chromosomes.

$$\begin{aligned} &\text{parents: } 011|010101|011; \quad 000|000111|111 \\ &\text{splitting: } 011 \quad 010101 \quad 011; 000 \quad 000111 \quad 111 \\ &\text{children: } 011|000111|011; \quad 000|010101|111 \end{aligned} \tag{4.3}$$

The method of choice for recombination is of course dependent on the genotype at hand. One can also reject recombination and just use mutation, similar to organisms that propagate by mitosis.

4.3.2 Mutation

Like in nature, mutation is a random change in the genotype. Again there are a multitude of ways of implementing mutation. One can use boolean mutation. There is a probability for a bit flip p, which is applied to the entire genome of bits of a new individual. Another

method is permutation, where parameters or bits in the genome switch position. If we mutate parameters α, not bits, the simplest way to mutate is to add a random variable corresponding to a previously set probability distribution, for example Gaussian mutation, Cauchy mutation or uniform mutation,

$$\alpha_j \leftarrow \alpha_j + \beta \qquad , \beta \propto \mathcal{N}(x_0, \sigma), \mathcal{C}(x_0, \gamma), \text{or } \mathcal{U}(a, b). \tag{4.4}$$

4.4 Example: Simple Regression Problem

As can be seen in the previous sections, one has a lot of choices to make if one wants to implement a genetic algorithm. To get a grip on the nature of genetic algorithms, we take a look at an example problem and how one could go towards solving it within this framework. For this purpose, we take a look at a simple quadratic regression problem: we want to minimize the mean square error. Hence, we choose our fitness function to be

$$f_i = -E^i_{ms} \tag{4.5}$$

("$-$" because we want the fitness to grow). (Of course one could also just minimize the error. We maximize the fitness in this example for conceptual reasons).
The individuals are polynomials of second degree,

$$I_i = a_i + b_i x + c_i x^2. \tag{4.6}$$

Thus the genes are the coefficients with the phenotype being the graph of the parabola. Each generation, we choose survivors with survival probability

$$p_i = \left(1 - \frac{-f_i}{\sum_{i=1}^{N} |f_i|}\right). \tag{4.7}$$

For a repopulation scheme, we use crossover and mutation. For crossover, we take two survivors and split the coefficients of each individual into two sets. The first offspring inherits the first coefficient set of the first parent and for the second offspring vice versa

$$\begin{aligned} &\text{parents: } a_1|b_1, c_1; \quad a_2|b_2, c_2 \\ &\text{splitting: } a_1 \quad b_1, c_1; \qquad a_2 \quad b_2, c_2 \\ &\text{children: } a_1|b_2, c_2; \qquad a_2|b_1, c_1. \end{aligned} \tag{4.8}$$

The children get mutated afterwards by randomly taking one of the coefficients and changing them by a Gaussian distributed number $\mathcal{N}(0, 0.1)$. This is done for $G = 2000$ generations. The result is summarized in Figure 4.1.

4.5 Advantages and Disadvantages of Genetic Algorithms

From the previous example, one can deduce a few advantages and disadvantages. Starting with the disadvantages [KCK20], the initial population has to be chosen a priori. If the population size is too large, it will significantly slow down the solution search. However if one chooses the population size too small, the solution will be poor. Usually, the population size

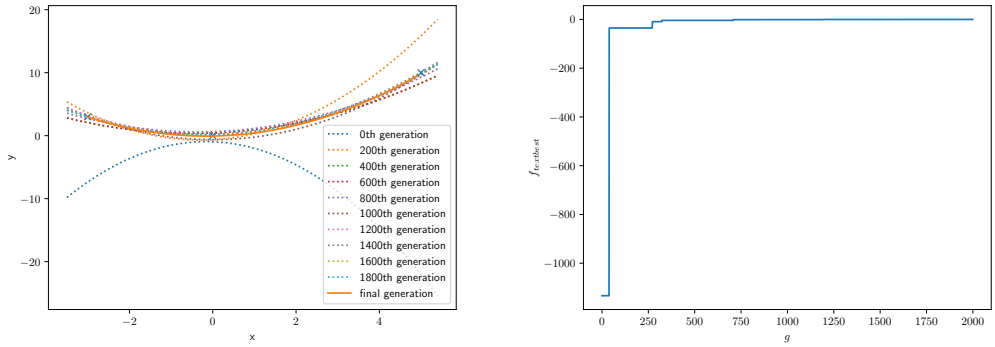


FIGURE 4.1: Solutions (l.) and fitness function (r.) of the polynomial regression with the given points (-3,3), (0,0) and (5,10). One can see that the described algorithm improves the result with the growing number of generations.

stays constant from generation to generation. There are methods with using variable population sizes [HL99; HH10], but their efficiency still relies on the initial choice.
Connected to a to small population size is premature convergence. This describes a loss in diversity of the population during the execution of the algorithm. This can happen due to a small population size but also due to the selection method used in the algorithm. It can happen, that the selection method is decreasing the diversity too much, leading to premature convergence.
Another issue is the fitness function. The fitness function is of course dependent on the problem at hand. However, finding a suitable fitness function is generally not a trivial task. This means, that one has to really understand the problem one wants to solve in depth.
Also, the degree between the crossover and mutation operators has to be kept in mind. If one completely disregards mutation, then the searchable solution space is limited by the initial population. However, the probabilistic nature of mutation makes an optimal ratio between the two operators difficult.
Another issue is the choice of encoding for the genome. There are are many possibilities, like binary (see eq.(4.3)) or value encoding (like was used in the prior example (section4.4)) but also octal, hexadecimal. More sophisticated encodings are the tree encoding, where the chromosomes are a tree of functions or commands [KCK20], and the permutation encoding, that is a string of numbers that represents a position in a sequence, that is mostly used in ordering problems. However, there is no general method of choosing an encoding scheme. Thus, an educated guess has to be made.
Another problem at the computational end is the speed and computational cost of genetic algorithms. These algorithms are usually quite slow. The example algorithm of the previous section is of course much slower than solving the regression problem with for example steepest descent. The execution can also be very computationally costly, especially with large population sizes and genomes.

However genetic algorithms have some advantages that makes them a valid choice for a solution finding algorithm.
One of the advantages is the possibility of parallelization. This can be implemented in various ways. If the selection method chosen is only dependent on the individual and not the entire population then this part can be parallelized. Mutation and crossover can be trivially

parallelized [San18].
If premature convergence is avoided, then a much larger part of the solution space is searched then in conventional solution finding algorithms. This also means that the chance of being stuck in a local optimum is way smaller than in other algorithms.
Genetic algorithms are also way better suited for highly complex and multi-objective problems due to the nature of the fitness function. As long as the problem can be feasibly be expressed in a fitness function, a problem can be solved with genetic algorithms. For multi-objective problems, another advantage is that the algorithm can give a multitude of solutions.
Due to the large solution space that is searched, genetic algorithms are also very resistant to noisy fitness functions.
They are also modular. This is useful, if one adapts to other problems, but also if one wants to use genetic algorithms in combination with other algorithms. For example, a current field of research is automated machine learning (AutoML). Deep learning, training artificial neural networks with back propagation schemes, is highly dependent on hyperparameters (learning rate, neural network architecture, choice of optimization scheme etc.). AutoML algorithms seek to find the optimal choice of hyperparameters. Some of them use genetic algorithms to find these optimal hyperparameter sets [VJ23].

Chapter 5

Stochastic Mechanics and Neural Networks

One of the disadvantages of stochastic mechanics is the non-trivial solution of the problem of finding the velocity fields u and v. In stochastic mechanics, they contain all of the information about a system, like the wave function in conventional quantum mechanics. The Bender/Steiner algorithm described in section 3.5 is not only very slow but also suffers heavily from the curse of dimensionality. In this chapter, a new algorithm will be introduced to solve this problem [HP23].
The algorithm will be a genetic algorithm (for general information, see chapter 4.5). Due to the modularity of genetic algorithms, we have to introduce many elements of genetic algorithms that had to be adapted to this problem.
Most results in this chapter are taken from "Machine learning quantum mechanical ground states based on stochastic mechanics" by Henk and Paul [HP23].

5.1 Derivation of the Fitness Function

We start with the quantum Hamilton principle [Pav95] introduced in eq. (3.4),

$$J[\hat{v},\hat{u}] = \min_{v}\max_{u} E\Big\{ \int \mathrm{d}t \Big[\frac{1}{2}m[v(x,t) - iu(x,t)]^2 - V(x,t)\Big] + \Phi_0 \Big\}. \tag{5.1}$$

In the following we will consider the stationary ground state only, i.e., the quantum mechanical action is given by $S(x,t) = -Et$, where E is the energy of the system, and limit ourselves to one-dimensional problems, which we will treat numerically later on. The above variation principle then simplifies to

$$J[u_0] = \max_{u} \mathbb{E}\Big\{ \int_0^T \mathrm{d}t \Big[-\frac{1}{2}mu^2(x(t)) - V(x(t))\Big] + S_0 \Big\}, \tag{5.2}$$

which is equivalent to

$$J[u_0] = \min_{u} \mathbb{E}\Big\{ \frac{1}{T}\int_0^T \mathrm{d}t \Big[\frac{1}{2}mu^2(x(t)) + V(x(t))\Big] + S_0 \Big\}, \tag{5.3}$$

The variation principle of eq. (5.3) aims to find an osmotic velocity such that the expectation value of the time-averaged energy along a particle trajectory is minimized. The integrand is the stochastic energy of the particle $H(x) = {}^{mu^2(x)}/_2 + V(x)$, which is a fluctuating quantity whose average is equal to the quantum mechanical expectation value of the Hamiltonian in

the ground state. Using the ergodicity of the Brownian diffusion in the ground state, this time-averaged energy is equal to the ensemble-averaged energy and we can write

$$J[u_0] = \min_u \int \mathrm{d}x \rho(x) \left[\frac{1}{2} m u^2(x) + V(x) \right] . \tag{5.4}$$

The quantum Hamilton principle of the stationary ground state is therefore equivalent to the Rayleigh-Ritz principle for the quantum mechanical ground state. The ground state osmotic velocity u_0 gives rise to a ground state density

$$\rho_0 = \exp \left\{ \frac{2m}{\hbar} \int^x u_0(x') \mathrm{d}x' \right\} . \tag{5.5}$$

The Rayleigh-Ritz principle is the variational principle in conventional quantum mechanics. To derive it, one starts with the expectation value for the energy

$$\langle \Psi | \hat{H} | \Psi \rangle . \tag{5.6}$$

Using $|n\rangle$ eigenfunctions of the Hamilton operator

$$\hat{H} |n\rangle = \varepsilon_n |n\rangle \tag{5.7}$$

and entering the a unit-matrix with a sum over the all projection operators

$$\langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi | \hat{H} \sum_n |n\rangle \langle n|\Psi\rangle = \sum_n \langle \Psi | \hat{H} |n\rangle \langle n|\Psi\rangle \tag{5.8}$$

$$= \sum_n \varepsilon_n | \langle \Psi|n\rangle |^2 \tag{5.9}$$

which gives

$$\sum_n \varepsilon_n | \langle \Psi|n\rangle |^2 \geq \varepsilon \sum_n | \langle \Psi|n\rangle |^2 = E_0 \langle \Psi|\Psi\rangle , \tag{5.10}$$

ergo

$$\langle \Psi | \hat{H} | \Psi \rangle \geq E_0 \langle \Psi|\Psi\rangle . \tag{5.11}$$

This, we can reformulate as a variational principle by parametrizing $|\Psi\rangle_\nu$, resulting in a minimization principle for the ground state energy

$$E_0 = \min_\nu \frac{\langle \Psi_\nu | \hat{H} | \Psi_\nu \rangle}{\langle \Psi_\nu | \Psi_\nu \rangle} \tag{5.12}$$

the Rayleigh-Ritz principle [Rit09]. The principle was developed independently from the Schrödinger equation and predates it by almost 20 years with polynomial parametrizations or Fourier series parametrizations in mind.

We derived the Rayleigh-Ritz principle for stochastic processes from the Pavon principle for ground states. If one starts from the standard Rayleigh-Ritz principle and uses that for the ground state

$$|\Psi\rangle = \sqrt{\rho} \tag{5.13}$$

holds, and the Hamilton operator for one-dimensional one-particle systems has the structure

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x), \tag{5.14}$$

one gets

$$E_0 = \min_{\rho} \int \sqrt{\rho(x)}\Big[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\Big]\sqrt{\rho(x)}\mathrm{d}x \tag{5.15}$$

$$= \min_{\rho} \int \left(\frac{\hbar^2}{2m}\frac{\rho'(x)^2}{4\rho(x)^2} + V(x)\right)\rho(x)\mathrm{d}x - \int \frac{\hbar^2}{2m}\left(\frac{\rho''(x)}{2}\right)\mathrm{d}x \tag{5.16}$$

which by entering the osmotic velocity

$$u(x,t) = \frac{\hbar}{2m}\frac{\rho(x)'}{\rho(x)} \tag{5.17}$$

becomes

$$E_0 = \min_{u} \int \left(\frac{1}{2}mu(x)^2 + V(x)\right)\rho(x)\mathrm{d}x - \int \frac{\hbar^2}{4m}\rho''(x)\mathrm{d}x. \tag{5.18}$$

One can see that one obtains an additional term compared to (5.4). This term is zero in the ground state due to the normalization condition

$$\int_{-\infty}^{\infty} \rho(x)\mathrm{d}x = 1 \tag{5.19}$$

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}^2}{\mathrm{d}x^2}\rho(x)\mathrm{d}x = 0. \tag{5.20}$$

Also the integrand of the second term can be rewritten using (5.5) as

$$\frac{\hbar^2}{4m}\rho''(x) = \left(mu^2 + \frac{\hbar}{2}u'\right)\rho \tag{5.21}$$

and with (5.18)

$$E_0 = \min_{u} \int \mathrm{d}x\left(\frac{1}{2}mu(x)^2 + V(x)\right)\rho(x) - \int \left(mu^2 + \frac{\hbar}{2}u'\right)\rho\mathrm{d}x \tag{5.22}$$

$$\tag{5.23}$$

or after rearranging

$$0 = \int \mathrm{d}x\left(\frac{1}{2}mu(x)^2 + V(x) - mu^2 + \frac{\hbar}{2}u' - E_0\right)\rho(x)\mathrm{d}x. \tag{5.24}$$

$$\tag{5.25}$$

The ground state probability density is nowhere zero, if

$$\frac{m}{2}u^2 + \frac{\hbar}{2}\frac{\mathrm{d}u}{\mathrm{d}x} = V(x) - E_0 \tag{5.26}$$

holds everywhere.
With E_0 given, eq. (5.26) is the Riccati equation,

$$y(x)' = f(x)y^2(x) + g(x)y(x) + h(x) \tag{5.27}$$

in its general form. If one takes the stationary Schrödinger equation

$$\left(\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right)\Psi(x) = E\Psi(x) \tag{5.28}$$

and transforms it with

$$\sigma = \frac{\Psi'}{\Psi}, \tag{5.29}$$

then one obtains

$$\sigma' + \sigma^2 - \frac{2m}{\hbar^2}(V(x) - E_0) = 0, \tag{5.30}$$

ergo the Ricatti equation with $y(x) = \sigma$, $f(x) = 1$, $g(x) = 0$ and $h(x) = \frac{2m}{\hbar^2}(V(x) - E_0)$ for the ground state of the problem, which is equivalent to the stationary Schrödinger equation (for more details, see Appendix D).
This results in two fitness functions that will be used in this algorithm

$$f_1 = \int \left(\frac{m}{2}u^2 + V\right)\rho\mathrm{d}x + \left|\frac{\hbar^2}{4m}\frac{\int \rho''\mathrm{d}x}{\int \rho\mathrm{d}x}\right| \to E_0 \tag{5.31}$$

The second uses Eq. (5.26) and E_0 from the first phase of the algorithm to improve the estimate for u_0

$$f_2 = \frac{m}{2}u^2 + \frac{\hbar}{2}\frac{\partial}{\partial x}u - V + E_0 \to 0. \tag{5.32}$$

We need two fitness functions, because of their specific properties. Since in f_1, both terms contain the probability density ρ, the osmotic velocity is well approximated in regions where ρ is large, but relatively poorly approximated otherwise. The reason is that the probability density acts like a weight function. For the second term, the absolute value has to be used, because elsewise the integral can become negative for some solutions. Negative values for the second term result in wrong solutions that would none the less minimize the fitness function. We also added the normalization condition in the denominator of the second term for better numerical stability.
f_2 has the problem, that E_0 is not a priori known but has the advantage, that ρ does not appear. Thus the osmotic velocity improves even in regions where the corresponding probability density values are small.
So we have two phases of the algorithm. The first gives a good first approximation of the osmotic velocity and the ground state energy. The second one improves the osmotic velocity in regions, where $\rho(x)$ is small. Now, to go back to chapter 4.5, we have to find a parametrization for the osmotic velocities that the individuals in the population will use.

5.2 Neural Networks

Artificial neural networks, also called neural networks, have become a multitool in all of the sciences. They are often used because they can approximate almost every function, only being limited by their architecture. Before discussing their use in this work, we have to clarify the terminology of neural networks.

5.2.1 Artificial Neuron

Neural networks owe their name to the artificial neurons (see Fig. 5.1), henceforth only called neurons, that they are made of. Neurons are based on perceptrons, a logic unit, introduced to simulate the nerve network in the brain. A perceptron is connected to multiple input values x_i and gives an output o_j by

$$o_j = \begin{cases} 1 & \sum_i w_{ij} x_i + b > 0 \\ 0 & \text{else} \end{cases} , \tag{5.33}$$

with w_{ij} being the weights, connecting each input x_i to the perceptron. The bias simulates the action potential of naturally neurons. Based on this, we get neurons

$$o_j = \phi \left(\sum_i w_{ij} x_i + b_j \right) \tag{5.34}$$

with ϕ being a so-called activation function.

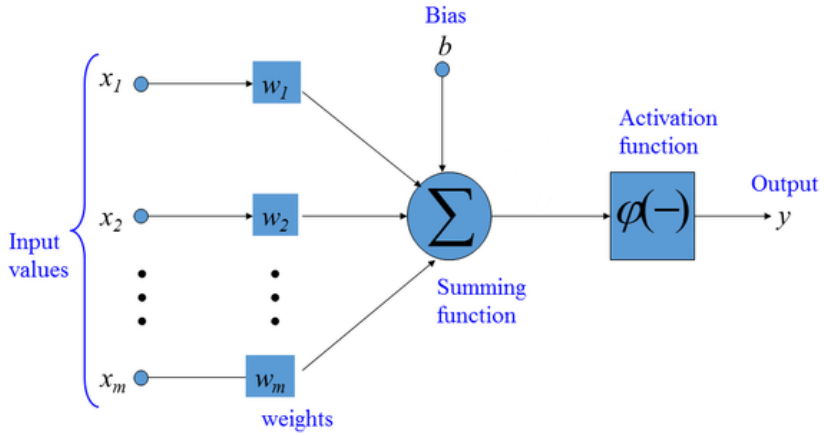


FIGURE 5.1: General scheme of a singular neuron [Alo+22] (*edited*). The input values x_i are each multiplied with the corresponding weight w_i and are summed over. Then the bias b is added. This value is then entered into the activation function $\varphi(\cdot)$, whose value is the output of the neuron.

5.2.2 Activation Function

The activation functions are usually non-linear, because the argument is already a linear combination of values. Moreover, if one wants to model a non-linear function, one needs non-linearity in the model. Activation functions help the neural network to approximate non-linear functions. If the approximated function is linear, the activation function can also

be omitted

$$\phi(x) = x. \tag{5.35}$$

Popular activation functions are the ReLU-function

$$\phi(x) = \begin{cases} x & x > 0 \\ 0 & \text{else} \end{cases}, \tag{5.36}$$

the positive definite sigmoid function

$$\phi(x) = \frac{1}{1 + e^{-x}}, \tag{5.37}$$

the hard limit function

$$\phi(x) = \begin{cases} 1 & x > 0 \\ 0 & \text{else} \end{cases}, \tag{5.38}$$

the softmax function

$$\phi(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}, \tag{5.39}$$

and the hyperbolic tangent

$$\phi(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{5.40}$$

The latter is used in this thesis due to it having positive and negative outputs, being non-linear and being easy to implement.
The choice of activation functions depends on the problem to be solved. For each layer, one could also have a specific activation function (see next section 5.2.3). Especially the activation function for the output layer should be handled with care, due to the limits an activation function puts on the possible output values.

5.2.3 Layers

A neural network is an ensemble of interconnected neurons. One has to choose how to interconnect a chosen number of neurons. For most function approximations, the arrangement of neurons into layers is a sensible choice. The two layers almost all neural networks have are the input layer and the output layer. As their name signals, the input values are put into the input layer and the output layer gives the output values. The hidden layers are in between those two layers. There are however many different kind of layers.
The most common one is the fully-connected layer, also called dense layer. In this kind of layer, every neuron is directly connected with each and every output of the previous layer and using these outputs as inputs for themselves. This kind of layer is found in almost all kinds of neural networks. They can become computationally costly, if they contain a large number of neurons. Neural networks that only consist of fully-connected layers are called feed-forward neural networks (see Fig. 5.2).
Another common type of layer are convolutional layers. These are very popular for image

classification tasks and perform a convolution between a kernel of weights and the input by way of the dot product. Neural networks that comprise of convolutional layers are called convolutional neural networks.
Recurrent layers function similar to a fully-connected layer. The neurons however calculate a hidden state in each step, that they enter into the next calculation, which makes them useful for time-dependent tasks by retaining information of prior states.
In deep learning, ergo training the neural network on previously collected data by using backpropagation to regress the weights and biases to fit the neural network to the data, it can be useful to employ dropout layers, which during the training phase set the weights an biases of a random number of neurons to zero, preventing the model to rely to heavy on single neurons. This procedure improves performance on new data, the neural network is not trained on.
The number of neurons in each layer, the number and kind of layers, and the activation function in each layer is called the architecture of the neural network.

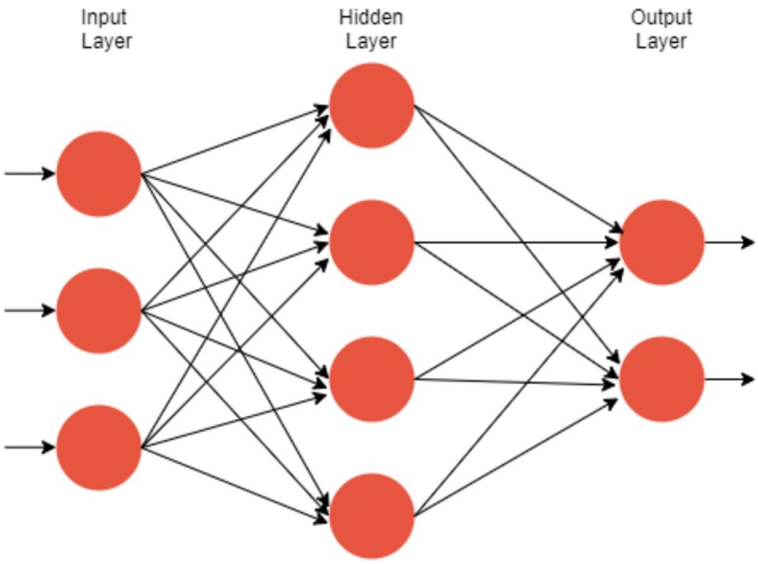


FIGURE 5.2: Example architecture of a feed-forward neural network [AGY23]. The arrowheads indicate where values are propagated to. The hidden and the output layer are fully-connected layer.

5.2.4 Neural Networks as (Almost) Universal Function Approximators

Neural networks were developed to solve the thirteenth of David Hilbert's 23 problems he presented on the International Congress of Mathematics in 1900 in Paris [Hil02]. The problem was formulated as follows: Is it possible to find a function that finds the solution to

$$x^7 + ax^3 + bx^2 + cx + 1 = 0 \tag{5.41}$$

as a chain of two-argument functions of the coefficients a, b, c? [Hil02]
Kolmogorov's theorem, also called Kolmogorov–Arnold representation theorem, states, that a three-layer neural network, meaning one input layer, one hidden layer and one output layer can approximate any multivariate continuous function, if the hidden layer has an infinite number

of neurons [Kůr92]. It was the first of what would later be called universal approximation theorems [Aug24]. Here, a few universal approximation theorems will be treated. A more exhaustive overview can be found in [Aug24] and [FD22].
Taylor's theorem (1715) states that any smooth and continuous function can be approximated at point a by the polynomial

$$f(x) = \sum_{k=0}^{\infty} \frac{(x-a)^k}{k!} f^{(k)}(a), \tag{5.42}$$

also called the Taylor series. Fourier's theorem (1807) says that any continuous and periodic function can be approximated by

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{N} \left(a_k \cos\left(\frac{2\pi k}{T} t\right) + b_k \sin\left(\frac{2\pi k}{T} t\right) \right), \tag{5.43}$$

with the Fourier series coefficients being

$$a_k = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) \cos\left(\frac{2\pi k}{T} t\right) \mathrm{d}t, \tag{5.44}$$

$$b_k = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) \sin\left(\frac{2\pi k}{T} t\right) \mathrm{d}t. \tag{5.45}$$

$$\tag{5.46}$$

It was later extended to non-periodic function, laying the foundation for the Fourier transform, discrete Fourier transform and the fast Fourier transform. Fourier's and Taylor's theorem can be called early universal approximation theorems.
The Weierstrass approximation theorem (1885) later stated that any function $f : [a, b] \to \mathbb{R}$ can be approximated with arbitrary accuracy ε by a finite polynomial

$$f(x) = \sum_{i=1}^{N} c_i x^i, \tag{5.47}$$

laying the foundations for polynomial regression [Aug24] and interpolation.
Kolmogorov's theorem (1959) states [Kol57], that any multivariate scalar function $f : [0, 1]^n \to \mathbb{R}$ can be written as

$$f(x_1, x_2, \ldots, x_n) = \sum_{j=1}^{2n+1} \beta_j \left(\sum_{n=1}^{n} \alpha_{ij}(x_i) \right), \tag{5.48}$$

with the functions $\alpha_{ij} : [1, 0] \to \mathbb{R}$ and $\beta_j : \mathbb{R} \to \mathbb{R}$.
With the popularization of neural networks many more universal approximation theorems were proposed for different activation functions, width (the number of neurons per layer) and depth (the number of layers). It has been shown [LF87] that any real-valued continuous function can be approximated with only two hidden layers and monotonic activation functions. In [IM88] it was proven that any arbitrary function can be approximated with a neural network with a single hidden layer with and infinite number of neurons. In [GW01] it was shown that

with one hidden layer and a monotonous "cosine squasher" activation function

$$\varphi(x) = \begin{cases} 0 & -\infty < x \leq -\frac{\pi}{2} \\ \frac{1}{2}(\cos\left(x + \frac{3\pi}{2}\right) + 1) & -\frac{\pi}{2} < x \leq +\frac{\pi}{2} \\ 1 & x \geq \frac{\pi}{2} \end{cases}, \tag{5.49}$$

the neural network can give the Fourier series approximation of any $L^2(\cdot)$ function on a compact set. All of this resulted in the theorem by Funashi, Hornick et al that any function function on the compact subset $f : \mathbb{X} \in \mathbb{R}^n \to \mathbb{R}$ can be approximated by the finite sum

$$\left| f(x) - \sum_{j=1}^{m} a_{2j}\sigma(\vec{a_{1j}}\vec{x} + b_{1j}) \right| < \varepsilon \qquad \forall x \in \mathbb{X} \tag{5.50}$$

with $\sigma(\cdot)$ being the sigmoid activation function (see eq. 5.37), meaning that any neural network with one hidden layer and the sigmoid approximation function can approximate any continuous multivariate function with arbitrary accuracy. Similar theorems exist for other activation functions [GI18], tanh, ReLU for univariate functions and one hidden layer. This architecture however can be proven to not be able to approximate multivariate functions with arbitrary accuracy [GI18]. Thus there are also univeral approximation theorems for arbitrary depth and fixed width.
In [Lu+17], it was proven that any Lebesgue-integrable function $f : \mathbb{R}^n \to \mathbb{R}$ can be approximated with a neural network with $n + 4$ or more number of neurons per layer and the ReLU activation function with arbitrary accuracy. For multivariate functions, it was proven in [Par+20], that that any Lebesgue-integrable function $f : \mathbb{R}^n \to \mathbb{R}^m$ can be approximated by a neural network with the width $\max(n + 1, m)$ with arbitrary accuracy.
There are many more universal approximation theorems. All of these show, that neural networks are a valid and solid choice of parametrization, if one has limited information about the solution.

5.2.5 Architecture for the Algorithm

Because we expect the osmotic velocity to be a continuous function, we choose an architecture with two hidden layers. For d-dimensional problems, the osmotic velocities are functions $u : \mathbb{R}^d \to \mathbb{R}^d$. Since we treat one-dimensional problems in this thesis, we only need one output and one input neuron. We chose tanh as activation function, because we expect the osmotic velocity to be smooth with positive and negative values. The osmotic velocity is not bounded between one and minus one, hence the output layer does not have an activation function. The number of neurons in the hidden layers chosen from experience to be $n = 7$ and $n = 25$ for each layer (Fig. 5.3).

5.3 Other Components and Overview of the Algorithm

Having found a parametrization of the osmotic velocity, more details of the algorithm should be discussed.

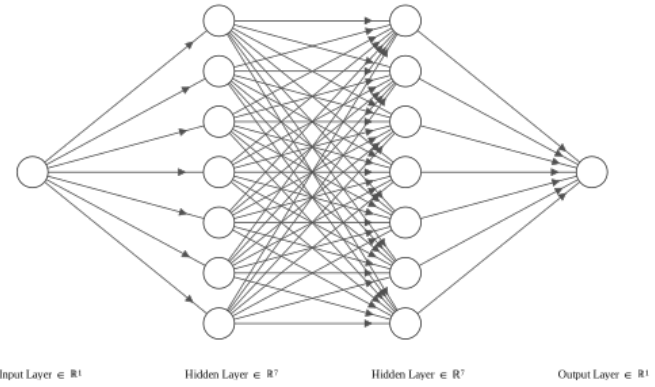


FIGURE 5.3: Architecture of the neural networks for the $n = 7$ case. The network consists of four layers, one input layer with one neuron, two hidden layer with n neurons and one output layer. Every layer except the output layer uses the tanh activation function. The output layer does not have an activation function.

5.3.1 Recombination and Mutation

Recombination is a non-trivial choice for neural networks. Montana and Davis [MD89] compared a few like crossover-weights, taking each weight and either putting in the value of the first or the second parent, and crossover-neurons, where instead of weights the entire neuron is picked from either the first or second parent. In [MD89], the operators were tested for a data classification problem. For our problem, recombination was abandoned to keep the algorithm simple and fast.
[MD89] also introduced mutation operators for neural networks. Firstly, unbiased-mutate-weights, having a fixed mutation probability for all weights and replacing it with a randomly distributed number. Biased-mutate-weights is having a fixed mutation probability for all weights and adding a random number to the unmutated weight. With mutate-neurons one takes a random number of neurons and bias-mutate every weight of that neuron. Mutate-weakest-neuron evaluates every neuron by setting for each neuron every weight to zero and comparing the output with and without the neuron being set to zero. The neuron with the smallest difference gets mutated. For simplicity, biased mutation was chosen. The added random number was Gaussian distributed $r \propto \alpha \cdot \mathcal{N}(0, 1)$, with α a set learning rate. The mutation only works if the genotype is a direct encoding, which was also chosen here.

5.3.2 Selection and Population Size

In chapter 4.5, a lot of different selection methods were described. Montana and Davis [MD89] calculated for each individual a survival probability depending on performance and selected randomly. Beeler et al. [Bee+19] however yielded good performances by simply picking the best performing neural networks. This method was chosen here. The population size was also fixed.

5.3.3 Workflow of the Algorithm

Having determined all details of the genetic algorithm, here is an overview of its workflow. The population size and the number of selected survivors were set by evaluation. This will be expanded on later in this chapter.
First, we have to initialize the population of 100 individuals. This is done by randomly picking for all neural networks in the initial population every weight and bias from a uniform distribution

$$w_{ij}^{p} \propto \mathcal{U}(-1,1) \qquad \forall p \in [1,100], i, j. \tag{5.51}$$

The following steps are repeated until convergence is reached. This is chosen to be having no improvement of the best performing network for ten generations.

- Calculate f_1 for each neural network
- Pick the 25 best neural networks, discard the rest
- Create 75 new neural networks by copying randomly from the pool of survivors and mutate these copies with Gaussian, biased mutation

One obtains a population of well-performing neural networks and the ground state energy E_0. The second phase is needed to get good results in the regions with small $\rho(x)$. Thus, we replace f_1 by f_2 and repeat until convergence is reached.

5.4 Calculating Ground States

The first task of the algorithm is to find the osmotic velocity of the ground state of three different one-dimensional potentials: the harmonic oscillator

$$V_{HO}(x) = \frac{\omega^2 m}{2} x^2, \tag{5.52}$$

the symmetric Pöschl-Teller potential

$$V_{PT}(x) = -\frac{\hbar^2}{2mx_0^2} \frac{l(l+1)}{\cosh(x/x_0)^2}, \tag{5.53}$$

and the quartic double well

$$V_{DW}(x) = \frac{V_0}{x_0^4}(x^2 - x_0^2)^2. \tag{5.54}$$

The harmonic oscillator is a standard problem in physics. The quantum harmonic oscillator is exactly solvable with equidistant energy eigenvalues

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \quad , n = 0, 1, 2..., \tag{5.55}$$

which lends it for testing numerical methods. In reduced units, the osmotic velocity for the harmonic oscillator is

$$u_{HO}(x) = -x. \tag{5.56}$$

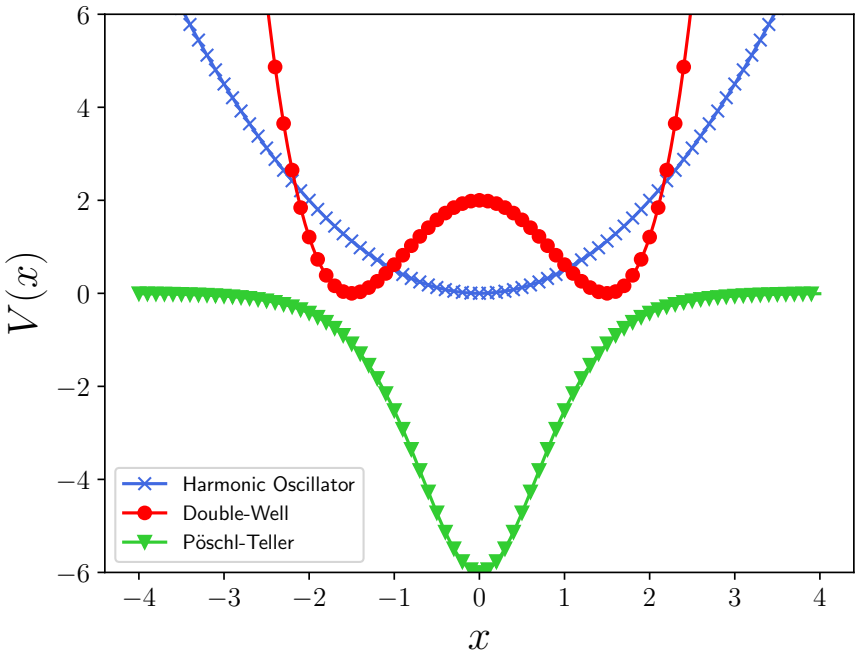


FIGURE 5.4: The harmonic oscillator,the Pöschl-Teller potential and the quartic double well potential in reduced units.

The symmetrical Pöschl-Teller potential, also called the modified Pöschl-Teller potential [Flü12], is used to describe two-atom systems, is also analytically solvable [PT33]. The wave functions are the Legendre-polynomials $P_l^\mu(\tanh(x))$ with l being the number of bound states and μ being the index of the bound state [Wil03]. The energy spectrum is given by

$$E_{PT}^{\mu} = -\frac{\hbar^2\mu^2}{2m} \quad \mu = 1, 2, ..., l-1, l. \tag{5.57}$$

For our calculations $l = 3$ was chosen.
The quartic double-well potential is only numerically solvable. The reference solution was created by solving the Schrödinger equation using the Numerov method [Bla67; Joh77] and using eq. (5.17) to obtain the osmotic velocity from the wave function.

5.4.1 Units

For the calculation, dimensionless units, also called reduced units, were used. For the harmonic oscillator the energy formulas yield

$$E_{HO} = \frac{m}{2}u(x)^2 + \frac{m\omega^2}{2}x^2, \tag{5.58}$$

$$\frac{E}{\hbar\omega} = \frac{1}{2}\frac{m}{\hbar\omega}u(x)^2 + \frac{1}{2}\frac{m\omega}{\hbar}x^2, \tag{5.59}$$

resulting for position, velocity and energy

$$\sqrt{\frac{m\omega}{\hbar}}x \to x \qquad \sqrt{\frac{m}{\hbar\omega}}u \to u \qquad \frac{E_{HO}}{\hbar\omega} \to E_{HO}. \tag{5.60}$$

For the Pöschl-Teller potential one arrives at

$$E_{PT} = \frac{m}{2}u(x)^2 - \frac{\hbar^2}{2mx_0^2}\frac{l(l+1)}{\cosh(x/x_0)^2}, \tag{5.61}$$

$$\frac{2mx_0^2}{\hbar^2}E_{PT} = \frac{1}{2}\frac{2m^2x_0^2}{\hbar^2}u(x)^2 - \frac{l(l+1)}{\cosh(x/x_0)^2}, \tag{5.62}$$

resulting for position, velocity and energy in

$$\frac{x}{x_0} \to x \qquad \frac{\sqrt{2}mx_0}{\hbar}u \to u \qquad \frac{2mx_0^2E_{PT}}{\hbar^2} \to E_{PT}. \tag{5.63}$$

The quartic double well potential does not have an obvious translation from reduced to SI-units. For our purposes

$$V_0 = 2 \qquad , x_0 = 1.5 \tag{5.64}$$

was chosen (see Fig. 5.4).

5.4.2 Osmotic Velocities, Probability Densities and Energies

Once the Algorithm is finished, one obtains the osmotic velocities and with that the probability densities and ground state energies of the three potentials.
The osmotic velocities are shown in Fig. 5.5. One can see, that there are no visual differences for large areas around $x = 0$. Slight derivations are visible at $|x| \geq 3.9$. In the subfigure of (5.5a), the result of the algorithm where the second phase with f_2 omitted is shown. There one can already see large differences at $|x| \geq 2.5$, hence the second phase of the algorithm improves the result significantly.
The probability densities were calculated with equation (5.5), shown in Figure 5.6. These results are more important, because the probability density is actually measurable, while the osmotic velocity is not. And due to the nature of equation (5.5), the slight errors in the osmotic velocity are not visible in the probability density, even in logarithmic scaling.
The ground state energy is also experimentally measurable and is thusly very important to be predicted correctly by any numerical algorithm. The energies can be observed in Table 6.1. These energies are predicted with high accuracy (small relative error). For consistency, the energies in that table are those calculated from the final osmotic velocity by

$$E_0 = \int \left(\frac{m}{2}u_0(x)^2 + V(x)\right)\rho_0(x)\mathrm{d}x \tag{5.65}$$

and not from the value calculated by the first phase using f_1. The value produced by the first phase differs only slightly from the second phase (in the sub-percentage scale).

Potential	correct E_0	calculated E_0	relative error
HO	0.5	0.50007	0.014%
PT	-4.5	-4.49939	0.0135%
DW	1.10342 [Kö+18]	1.10337	0.00453%

TABLE 5.1: Correct and calculated ground state energies for the three potentials with relative error in reduced units.

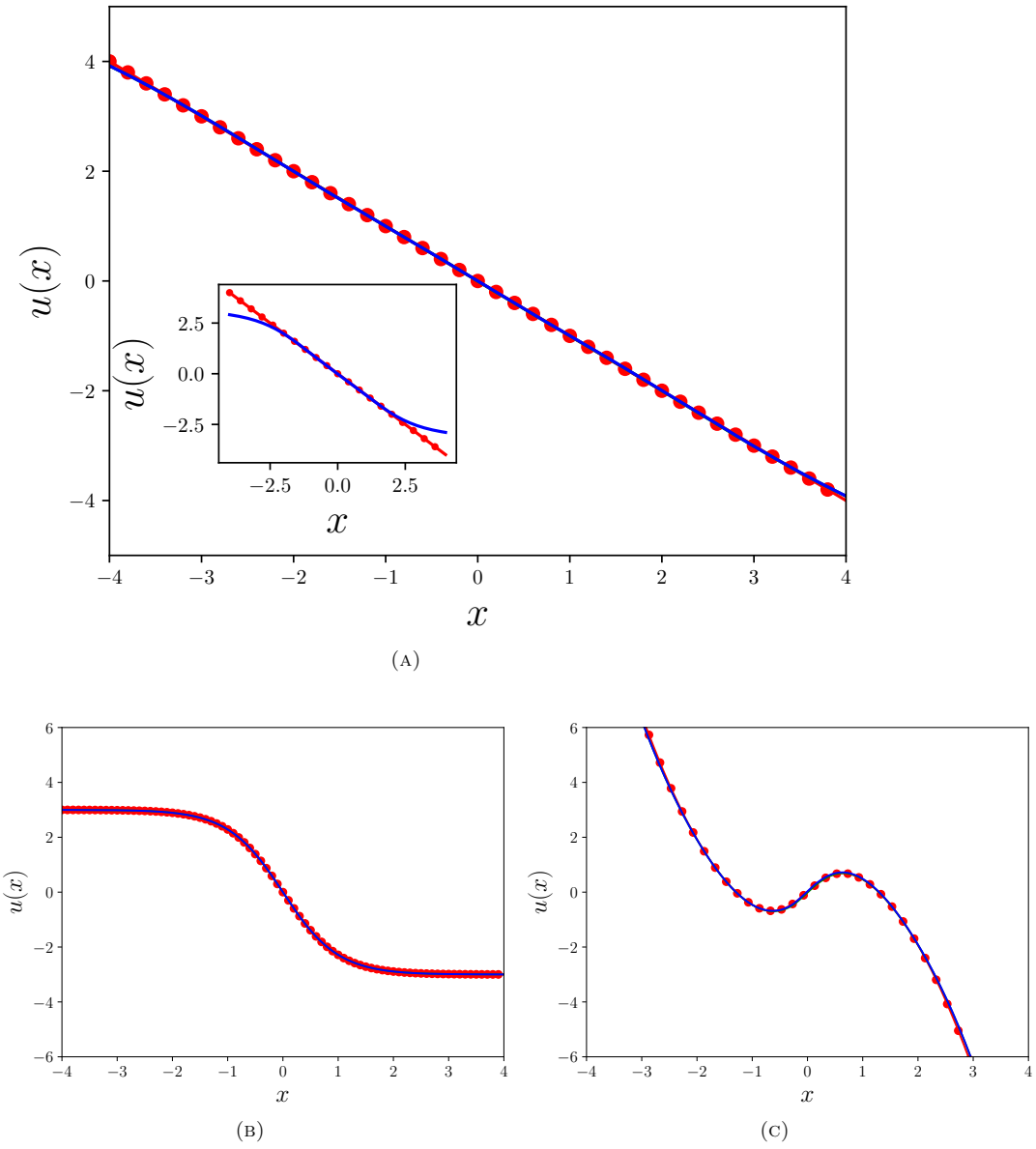


FIGURE 5.5: Osmotic velocities of the harmonic oscillator (top), Pöschl-Teller potential (bottom left) and quartic double well (bottom right) in reduced units. Visually, there is no difference between the references (connected red dots) and the solutions of the algorithm (continuous blue lines). In the subplot of (A) one can see the result if the second phase of the algorithm is omitted. There, the difference is much larger and clearly visible.

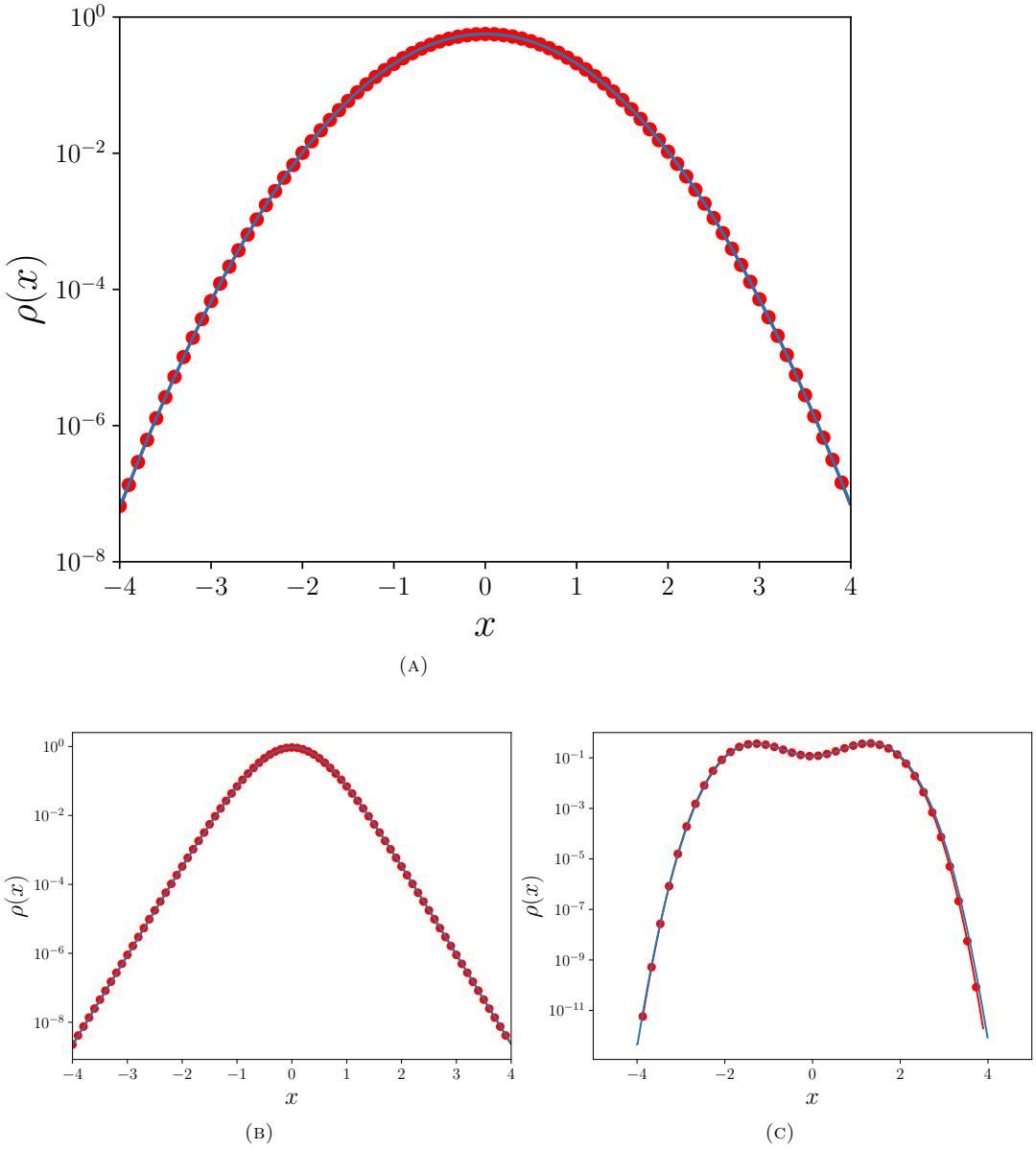


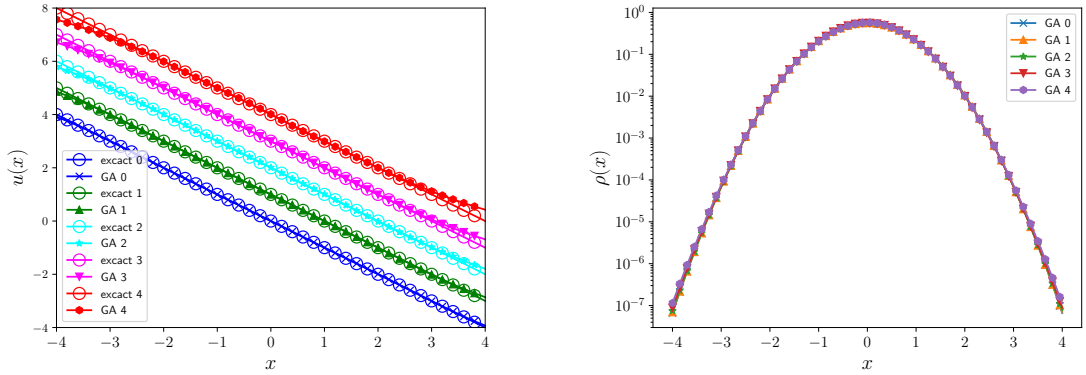
FIGURE 5.6: Probability densities of the harmonic oscillator (top), Pöschl-Teller potential (bottom left) and quartic double well (bottom right). All figures use reduced units. No visible differences between the references(connected red dots) and the corresponding probability densities to the calculated osmotic velocities (continuous blue line) (eq. 5.5) exist.

5.5 Exited States

As shown in section 3.6, the super-symmetric partner Hamiltonians are given by

$$\hat{H}_n = T + V_0 - \sum_{i=0}^{n-1} \frac{\partial}{\partial x} u_i. \tag{5.66}$$

Because the osmotic velocities in the sum are not known a priori, one has to solve the ground state and the exited states until the $n-1$-th state, iteratively until one obtains the n-th partner Hamiltonian. Thus one has to repeat the algorithm n times for each partner Hamiltonian to



(A) (B)

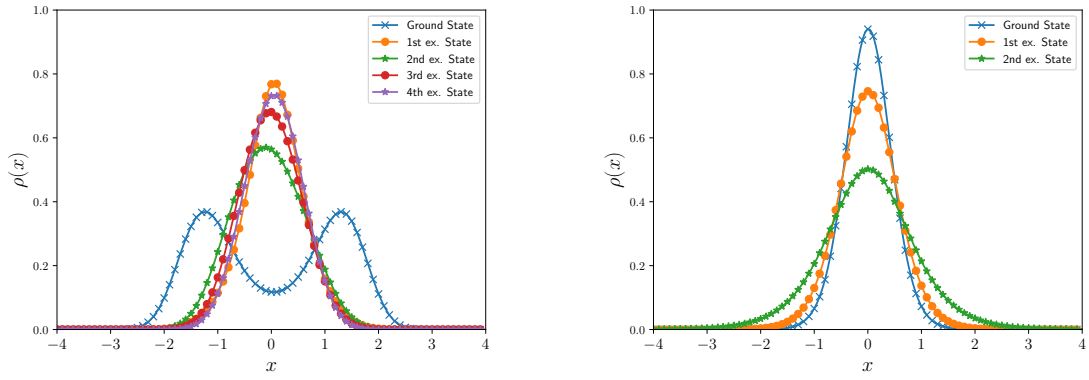
FIGURE 5.7: Osmotic velocities for the first four exited states and the ground state of the harmonic oscillator, separated by ΔE (for visibility), (left) and corresponding probability densities (right)(given by eq. 5.66) (GA n indicates the result of the genetic algorithm for the ground state $n = 0$ and the nth exited state). The error in the osmotic velocity increases with excitation. This has only very small influence on the error in the corresponding probability densities.

get the exited states.
These are however not entirely the correct exited states. Exited states wave functions have nodes. There the probability distribution becomes zero and because of eq. (5.17), there are singularities in the osmotic velocity. The ground states of the SUSY partner Hamiltonians however are nodeless.

As can be seen in Figures 5.7 and 5.8, the osmotic velocities are approximated very well. At $x \geq 3.5$ the slight numerical errors of the lower states are propagated by each excitation cumulate towards larger, but still small deviations from the reference values. The harmonic oscillator is a suitable potential, where this can be observed, because the Hamiltonian only changes by a constant

$$-\frac{\partial}{\partial x}u(x) = -\frac{\partial}{\partial x}(-x) = 1 = \Delta E \tag{5.67}$$

between exited states. The osmotic velocity is left unchanged. In Fig. 5.7 one can see the growing error in the osmotic velocity very well. But the error in the corresponding probability densities (unchanged from state to state) is very small. The probability density does not change from excitation to excitation, because we calculate the ground state of the supersymmetric partner Hamiltonians and not the true exited states. This propagated error can also be observed in Fig. 5.8 as the slight rightwards shift of the center of the probability densities of the higher exited states, which should be at $x = 0$ because of the symmetry of the potential.
Three exited states were calculated for the Pöschl-Teller potential, because we chose $l = 3$, the number of bound states. If one would try to calculate the 4-th exited state, the algorithm would give wrong results: The state corresponding to the higher energy would not be a bound state but a scattering state. Because the Rayleigh-Ritz-principle is the basis for this algorithm, the algorithm, like the Rayleigh-Ritz principle does not work for scattering states [Bur77]. The harmonic oscillator and the quartic double well do not have that problem, because

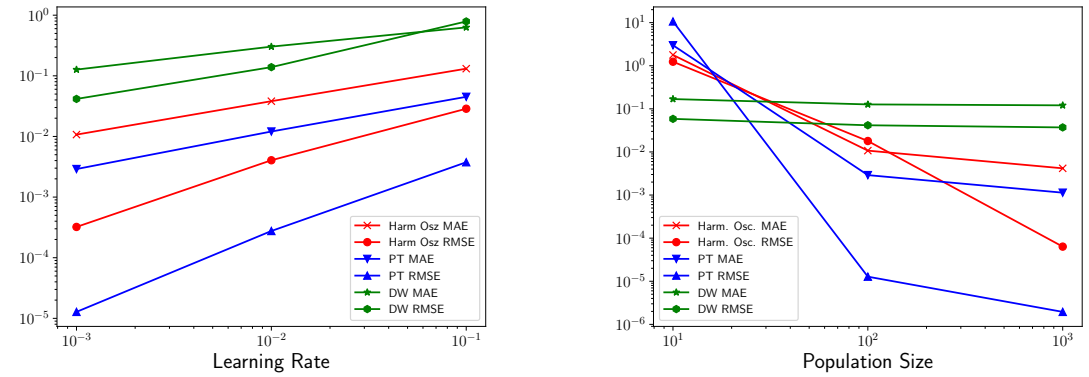


(A) (B)

FIGURE 5.8: Densities for the first four exited states and the ground state for the quartic double well (left) and the first two exited states and the ground state of the Pöschl-Teller potential (right). The error of the numerical calculation of the divergence in eq. (5.66) propagates and grows from excitation to excitation resulting in a slight shift of the symmetry axis of the densities (should by at $x = 0$).

$\lim_{x\to\pm\infty} V(x) = \infty$ and thus these are no scattering states.

5.6 Discussion of Hyperparameters and Parameters



(A) (B)

FIGURE 5.9: RMSE and MAE for the different potentials versus learning rate (left) and population size (right).

The algorithm contains a set of hyperparameters which can be adjusted to optimize the performance of the algorithm. These are called hyperparameters to differentiate them from parameters which are the weights and biases in the neural networks. Typically, the learning rate α should be small, to allow for a more accurate, sensitive search of the solution space and the population size should be rather large to allow for search in a larger volume of the solution space.

For evaluation purposes, the the relative energy error

$$\Delta E = \frac{|E_0 - E_{0\text{reference}}|}{E_{0\text{reference}}}, \tag{5.68}$$

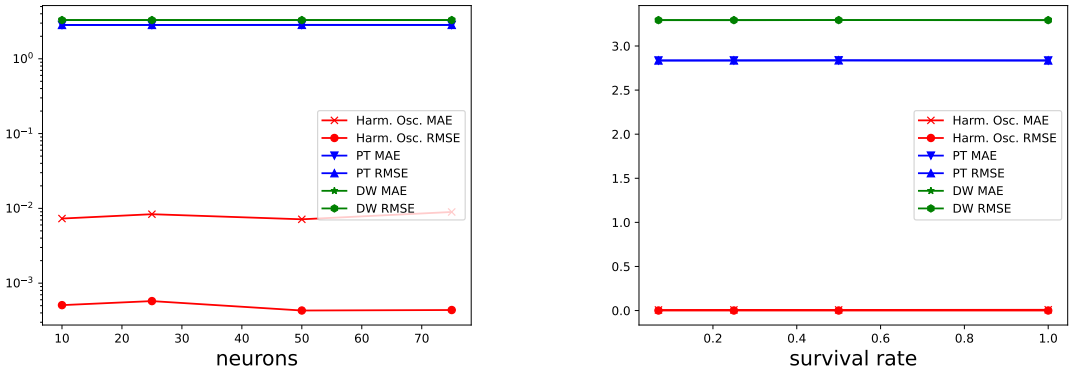
(N the number of discrete positions) mean absolute error MAE

$$\Delta_{MAE} = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{x_i}|, \tag{5.69}$$

and root mean square error RMSE

$$\Delta_{RMSE} = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(x_i - \hat{x_i})^2} \tag{5.70}$$

with respect to the reference curves were used. A learning rate of $\alpha = 0.001$ was found most useful (see Fig. 5.10a). Larger values resulted in large deviations in the RMSE and MAE were found (see Fig. 5.10a) with a decrease in magnitude in the errors for a decrease in magnitude in α. Experience shows that the population size should not be smaller than 100. Larger populations improve the results, but an increase in order of magnitude did not have the same magnitude in improvement in the errors (Fig. 5.10b). The 25% survivor rate was chosen, because an increase didn't result in a increase in accuracy but lowering the rate did result in lower accuracy. So at least for the problems described here, 25% were optimal.
Regarding the architecture, two hidden layers were sufficient for these problems with no significant improvement by a third hidden layer. This is consistent with the corresponding universal approximation theorem for real-valued continuous functions [LF87].
For the number of neurons n, experience showed that for the harmonic oscillator and the Pöschl-Teller potential $n = 7$ was adequate. For the double well potential, $n = 25$ was used (see Fig. 5.3)



(A) (B)

FIGURE 5.10: RMSE and MAE for the three potentials versus number of neurons (left) and survival rate (right).

5.7 Advantages over the Bender/Steiner Algorithm

Being a genetic algorithm, the proposed algorithm has all the advantages and disadvantages already discussed in section 4.5. The here described genetic algorithm has a many advantages over the stochastic optimal control algorithm described in section 3.5.
Firstly, it is conceptually easier to understand an energy minimization scheme than a scheme with two coupled forward backward stochastic differential equations. It is thusly also easier

to program and the program is less prone to bugs. The algorithm is also much faster and yields a much better approximation in a much larger region of the coordinate space.

Chapter 6

Time Series Analysis and Phase Space Dynamics

One of the big advantages of the stochastic mechanics approach to quantum mechanics is the treatment of time. In conventional quantum mechanics, due to the absence of a time-measurement operator, it is difficult to make predictions on, for example, how long a process takes on average. In stochastic mechanics, this is not a problem: One solves the stochastic differential equations of the problem and measures the time it takes for each realization of the process and takes the average.
Having stochastic trajectories of quantum processes has also the advantage of having a signal over time. Thus, one can use all the tools from time series analysis. Also, one can derive phase space portraits of quantum states.
Most of the results can also be found in this preprint *Levitodynamics: An analysis of quantum fluctuations based on stochastic mechanics*, Henk and Paul [HP25].

6.1 Introduction to Time Series Analysis

Time series analysis is a set of methods concerning time series, a set of data x_t measured at a specific time t. The two main types of time series are discrete and continuous time series, with only the former being treated in this thesis. The aim of time series analysis is to obtain additional insights from the time series. But first, we have to check what kind of time series the stochastic trajectories of quantum systems are.

6.2 Stationary and Ergodic Time Series

In general, stationary time series are such that their key properties do not change over time [BD16]. In physics, stationary generally means that the (mean) energy is constant and that pair correlations only depend on the time difference, like is the case in quantum mechanical ground states. In time series analysis however, one distinguishes different kinds of stationary systems.

Definition 2 *$\{X_t\}$ is a strictly stationary time series if two subsets $(X_1, \cdots, X_n)$ and $(X_{1+h}, \ldots, X_{n+h})$ share the same joint distribution function for all integers $h, n \geq 1$ (noted with $(X_1, \cdots, X_n) \stackrel{d}{=} (X_{1+h}, \ldots, X_{n+h})$) [BD16].*

This is definitely the case for the trajectories of quantum mechanical ground states, because the X_t values are proportional to a probability distribution, $X_t \propto \rho(x)$, that does not change with time.

Definition 3 *$\{X_t\}$ is a weakly stationary (sometimes called wide sense stationary) time series if*

- *the mean of the time series is independent of time,*
- *the covariance function is $\gamma_X(t+h,t)$ is independent of time t*

[BD16].

Before the covariance function is treated, we should take a look at some properties of strictly stationary time series [BD16]:

- X_t are identically distributed,
- $(X_1, \cdots, X_n) \stackrel{d}{=} (X_{1+h}, \ldots, X_{n+h})$,
- $\{X_t\}$ is weakly stationary if $\mathbb{E}[X_t^2] < \infty$ for all times t.

With the last property, we can show that the stochastic trajectories in this work are weakly stationary. The Heisenberg uncertainty relation is fixed for quantum mechanical ground states ergo

$$\Delta x \Delta p \geq \frac{\hbar}{2} = \text{const.} = c \tag{6.1}$$

which yields

$$\Delta x = \sqrt{\mathbb{E}[x^2] - \mathbb{E}[x]^2} \tag{6.2}$$

$$c = (\mathbb{E}[x^2] - \mathbb{E}[x]^2)\Delta p \tag{6.3}$$

$$\mathbb{E}[x^2] = \frac{c}{\Delta p} + \mathbb{E}[x]^2 < \infty \tag{6.4}$$

for $\Delta p \neq 0$ and $\mathbb{E}[x] < \infty$, which is the case for physical systems. This means that they are weakly stationary.
Ergodicity is another important property of time series and concerns the moments.

Definition 4 *If the time average of a process is equal to the mean from the probability distribution*

$$\frac{1}{T}\int_0^T x_t \mathrm{d}t = \int x\rho(x)\mathrm{d}x \tag{6.5}$$

for large enough times T, the process is mean-ergodic or ergodic in the first order.

Closely connected to stationarity and ergodicity are the autocovariance and autocorrelation functions [BD86].

Definition 5 *The autocovariance function (ACVF) is defined via*

$$\gamma_{XX}(s,t) = Cov(X_s, X_t) = \mathbb{E}[(X_s - \mathbb{E}[X])(X_t - \mathbb{E}[X])], \tag{6.6}$$

with the autocorrelation function (ACF) being

$$r_{XX}(s,t) = \frac{\gamma_{XX(s,t)}}{\gamma_{XX(0,0)}}. \tag{6.7}$$

For weakly stationary processes, both functions depend only on the time difference $\tau = s - t$:

$$\gamma_{XX}(\tau) = \mathrm{Cov}(X_{t+\tau}, X_t) = \mathbb{E}[(X_{t+\tau} - \mathbb{E}[X])(X_t - \mathbb{E}[X])], \tag{6.8}$$

$$r_{XX}(\tau) = \frac{\gamma_{XX}(\tau)}{\gamma_{XX}(0)}. \tag{6.9}$$

The ACVF and ACF are useful since they measure the degree of interdependence between the values of a time series [BD16].

Definition 6 *Processes are second-order ergodic, if*

$$\lim_{T\to\infty} \mathbb{E}[\langle r_{XX}(\tau)\rangle_T] = r_{XX}(\tau), \tag{6.10}$$

$$\lim_{T\to\infty} Var[\langle r_{XX}(\tau)\rangle_T] = 0. \tag{6.11}$$

Second-order ergodic process are sometimes called wide sense ergodic. Processes are ergodic, if one can deduce ensemble statistics from single sample paths, if the paths are long enough. This is the case for the stochastic trajectories of quantum mechanical ground states, which are thus ergodic in the first order. All ergodic processes are stationary, however not all stationary processes are ergodic (for further discussion, see Appendix C.3).

6.3 Sampling the Ground States

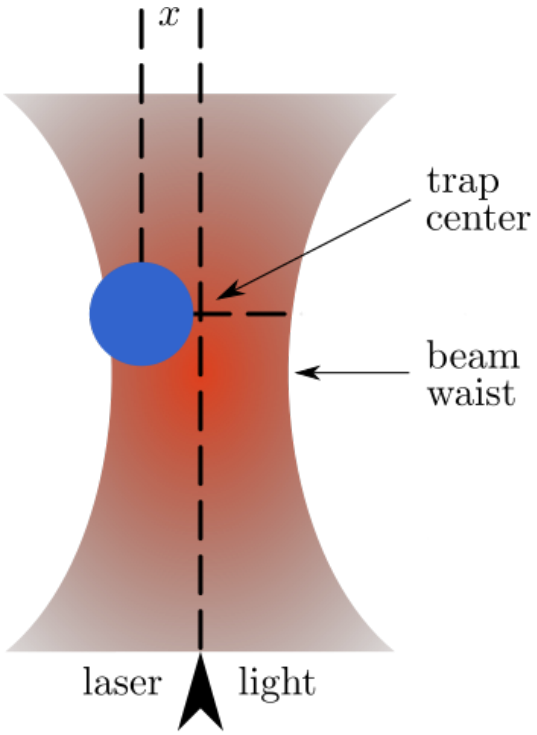


FIGURE 6.1: Scheme of an optical tweezer (taken from Wikimedia Commons [Com24](*edited*)). A dielectric particle is trapped in the intensity maximum due to the force acting along the gradient of the electric field of the laser.

This analysis in this chapter was inspired by the work of Magrini et al. [Mag+21]. In the experiment [Mag+21], a dielectric silica sphere of diameter 70 nm with a mass of $m \approx 2.8\,10^{-18}$ kg is trapped in the intensity maximum of the focus of a laser beam (see Fig. 6.1). The form of

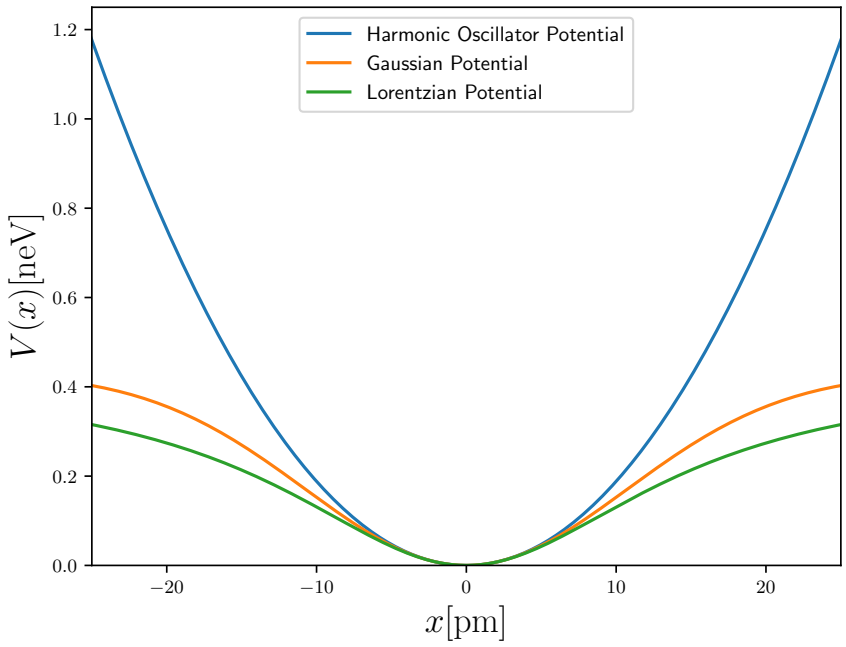


FIGURE 6.2: The three potentials, that were analyzed. All of them approximate the harmonic oscillator in the first order.

the confining potential is given by the shape of the laser beam which is often approximated as Gaussian (we call the relevant position variable x in the following)

$$V_G(x) = V_0 \left(1 - e^{-\frac{x^2}{2x_0^2}}\right) \tag{6.12}$$

or Lorentzian

$$V_L(x) = V_0 \left(1 - \frac{1}{\frac{x^2}{2x_0^2} + 1}\right) . \tag{6.13}$$

Both can be approximated by a harmonic potential

$$V(x) = \frac{m\omega^2}{2} x^2 = \frac{V_0}{2} \frac{x^2}{x_0^2} \tag{6.14}$$

(see Fig. 6.2). The energy and length scales are set by V_0 and x_0, respectively. The typical length scale for the harmonic oscillator is $x_0^2 = \frac{2\hbar}{m\omega}$ with the energy scale $V_0 = \hbar\omega$. The three potentials can be seen in Fig. 6.2. The Gauss and Lorentz potentials are in first order approximation equal to the harmonic oscillator. However both have a softer confinement of the particle than the harmonic oscillator. The ground state energies differ as well (see Tab. 6.1).

Potential	E_0
harm. Osc.	0.50004
Gaussian	0.40559
Lorentzian	0.36484

TABLE 6.1: Ground state energies of the three potentials.

Using the algorithm from chapter 5.7, we obtain the osmotic velocities of the three potentials depicted in Fig. 6.3. Using the Heun-scheme (see section 2.3), we solve the stochastic differential equation with discrete time step $\Delta t = 10^{-4}$ for 10^7 time steps, obtaining three sample paths (see Fig. 6.4),that are to be treated with time series analysis. In [Mag+21], only harmonic confinement is considered. Even though all potentials simplify to the harmonic oscillator in the first order, one has to inspect the possible differences important to experiments.

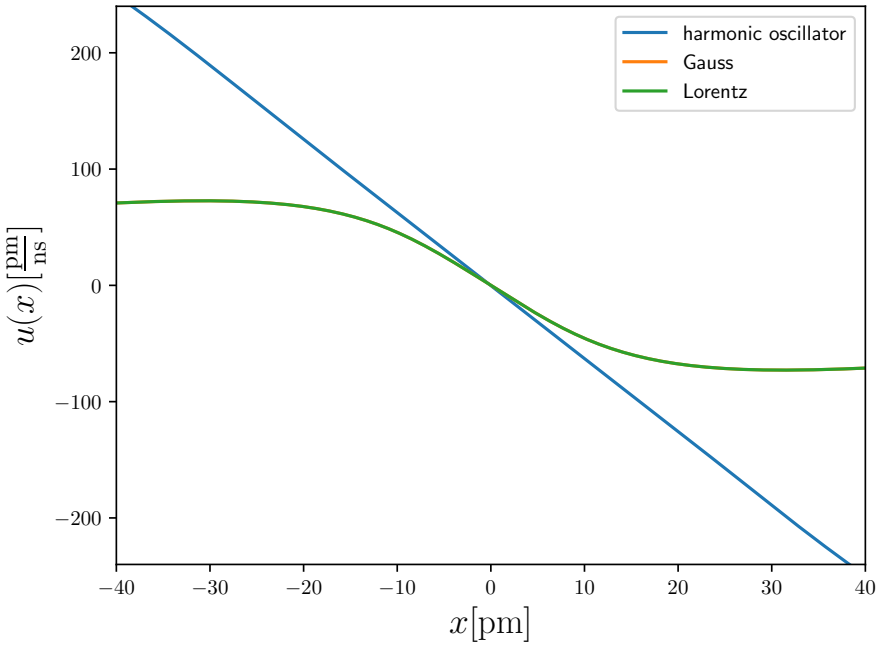


FIGURE 6.3: Osmotic velocities of the three potentials. The Gauss and Lorentz velocities are almost identical, even though their potentials differ significantly.

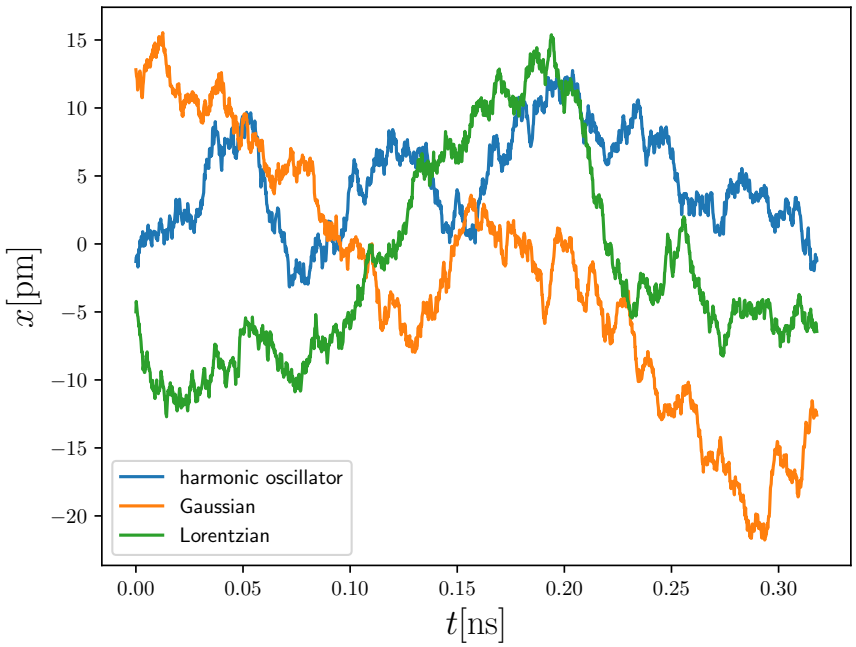


FIGURE 6.4: Three sample paths to be treated with time series analysis. There are no immediate, qualitative differences observable, making it necessary to treat them with statistical methods.

6.4 The Autocorrelation Function

Correlation is a measure of statistical dependence between two observations. This statistical dependence does not have to be causal. It is defined by

$$r_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\mathbb{E}[(X - \langle X \rangle)(Y - \langle Y \rangle)]}{\sigma_X \sigma_Y}, \tag{6.15}$$

with $\sigma_{X/Y}$ being the standard deviations of the two processes, introduced to normalize the corelation to the interval $[-1, +1]$ which is not the case for the covariance. If the correlation is at $+1$, the two variables X and Y are fully correlated,at -1 fully anti-correlated and at 0 uncorrelated.
ACF shows how much a signal is correlated with itself over a time difference τ, which does not depend on the reference time for stationary problems. The ACF starts at $r_{XX}(\tau = 0) = 1$ and converges to $\lim_{\tau \to \infty} r_{XX}(\tau) = 0$ per the above definition.

To calculate the ACF for the sample paths of the three potentials, each sample path was separated into windows which ended all at the last point of the sample path. The starting point varied. Here, every 1000th value was used as starting point. For each window, the ACF was calculated and finally over each window-ACF was averaged.
There are slight differences observable in the ACFs of the three potentials (see Fig. 6.5). All show the expected behavior of starting at 1 at $\tau = 0$ and converging to 0, ergo uncorrelation for $t \to \infty$. There is an area of $r_{XX}(\tau) < 0$ due to numerical errors. The decay of the harmonic oscillator ACF is faster, ergo it decorrelates more quickly than the other two. The Gauss and Lorentz ACFs are quite similar to one another.

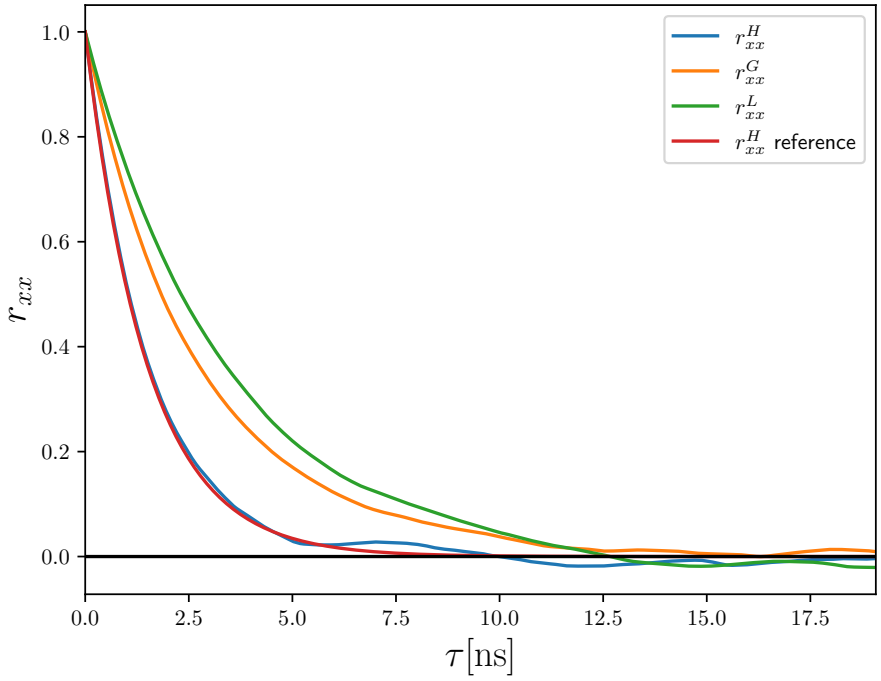


FIGURE 6.5: ACF of the three potentials. The three lines only differ marginally. The reference ACF of the harmonic oscillator is the analytic result for the Ornstein-Uhlenbeck process.

All three oscillators show however very similar behavior in the ACF although their potentials differ substantially.

6.5 The Power Spectral Density

The power spectral density (PSD), also called power spectrum, is a measure of how much of the total power in a signal comes from each frequency. It can be used in an experimental setting to find sources of and reduce noise in the signal.
The PSD is defined as

$$S_{XX}(f) = \lim_{T\to\infty} \frac{1}{T}|\mathcal{F}(X_t)|^2, \tag{6.16}$$

where $\mathcal{F}(X_t)$ is the Fourier transformed signal.
Using the convolution theorem

$$\lim_{T\to\infty} \frac{1}{T}|\mathcal{F}(X_t)|^2 = \lim_{T\to\infty} \frac{1}{T}\mathcal{F}(X_t)\mathcal{F}(X_t) = \lim_{T\to\infty} \frac{1}{T}\mathcal{F}(X_t * X_t) \tag{6.17}$$

$$= \int \left[\lim_{T\to\infty} \frac{1}{T}\int X_t^*(t-\tau)X_t(t)\mathrm{d}t\right] e^{-i2\pi f\tau}\mathrm{d}\tau, \tag{6.18}$$

tells that the inner integral is equal to the autocorrelation function, if the signal X_t is ergodic. Thus, we arrive at the Wiener-Khinchin theorem [Khi34], that the PSD is the Fourier

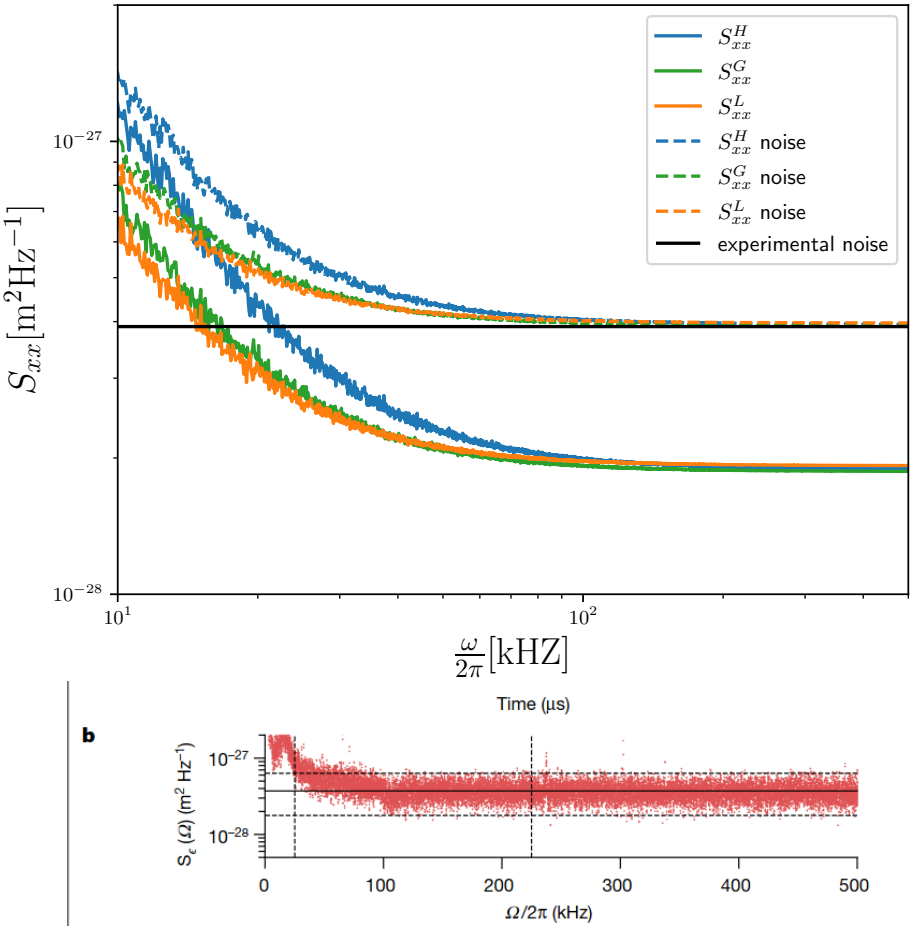


FIGURE 6.6: PSD of the three potentials with and without added experimental noise (top) to compare with experimental data published [Mag+21] (bottom). The slight differences in the theoretical PSD would not be visible in an experimental setting due to experimental noise.

transform of the autocorrelation function

$$S_{xx}(f) = \int_{-\infty}^{\infty} r_{xx}(\tau) e^{-i2\pi f\tau} \mathrm{d}\tau = \mathcal{F}(r_{XX}(\tau)). \tag{6.19}$$

This gives us two methods for obtaining the PSD, either by Fourier transforming the signal directly or by Fourier transforming the ACF. Practically, if one would use the former method the use of Welch's method [Wel67], that is segmenting the signal into overlapping windows, Fourier transforming those (these Fourier transformed signals are called periodigrams) and averaging over them, is preferable.

In this thesis, we use the Wiener-Khinchin theorem, obtaining the PSD $S_{xx}(f)$ by Fourier transforming the ACF from the previous section resulting for real-valued autocorrelation functions in

$$S_{xx}(f) = \int_{-\infty}^{\infty} r_{XX}(\tau) \cos(2\pi f\tau) \mathrm{d}\tau = \mathcal{F}[r_{XX}(\tau)], \tag{6.20}$$

resulting in Fig. 6.6. The results are consistent with the experimental results from [Mag+21], with a decay into experimental noise to 100kHz. Again, the results for the Lorentz and Gauss potential are quite similar to each other, with the small differences from the ACFs propagating to the PSDs. The decay of the harmonic oscillator into the residual experimental noise is slower compared to the other two. However, the residual noise in the experiment [Mag+21] should be to large for the differences in the PSDs to be measurable.

6.5.1 Units

For simulation purposes, we used reduced units as before (section 5.4.1)

$$\frac{x}{x_0} \to x \qquad \frac{\mathrm{d}x}{x_0} \to \mathrm{d}x \tag{6.21}$$

The potentials have a generic V_0 prefactor which for the harmonic oscillator is equal to $\hbar\omega$. Because the other potentials originate from laser potentials, they depend on the strength of the laser, ergo

$$V_0 \propto |\vec{E}^2|. \tag{6.22}$$

Starting from the energy, reduced by using $\hbar\omega$

$$\frac{E}{\hbar\omega} = \frac{p^2}{2m\hbar\omega} + \frac{m\omega x^2}{2\hbar\omega} \tag{6.23}$$

resulting in position and momentum

$$x_0 = \sqrt{\frac{\hbar}{m\omega}}, \tag{6.24}$$

$$p_0 = \sqrt{m\hbar\omega}, \tag{6.25}$$

respectively. For the time, we use

$$t_0 = \omega^{-1}. \tag{6.26}$$

To facilitate the comparability with the experiment [Mag+21], we put in the settings from the experiment

$$m \approx 2.8 \cdot 10^{-18}\mathrm{kg}, \tag{6.27}$$

$$\omega = 2\pi \cdot 10^4\mathrm{kHz}, \tag{6.28}$$

to obtain SI units for the simulations:

$$t_0 = 0.159 \cdot 10^{-9}s, \tag{6.29}$$

$$x_0 = 7.742 \cdot 10^{-13}m, \tag{6.30}$$

$$p_0 = 1.884 \cdot 10^{-19}\mathrm{kg}\frac{\mathrm{m}}{\mathrm{s}}, \tag{6.31}$$

$$u_0 = \frac{p_0}{m} = 67.3\frac{\mathrm{pm}}{\mathrm{ms}}. \tag{6.32}$$

These are used in Fig. 6.5 and 6.6. For the analysis of phase space dynamics, we use reduced units.

6.6 Phase Space Dynamics

As shown in [Mag+21], it is possible to extract a phase portrait of a quantum particle from an experimental setup. Therefore, it is desirable to deduce a theoretical method for obtaining such trajectories. For classical systems, this is a simple task. But for quantum systems one has to find phase space trajectories, which obey the Heisenberg uncertainty principle.
The Heisenberg uncertainty relation is one of the foundational principles of quantum mechanics. It was motivated in 1927 using the γ-microscope thought experiment (see section 3.2 for a detailed discussion). Its connection to Schrödinger's wave mechanics [Sch26b; Sch26c; Sch26d; Sch26e] is evident, if one considers the Fourier transformation [Qui06] of the wave function in position space, i.e. the wave function in momentum representation

$$\tilde{\Psi}(p) = \mathcal{F}[\Psi(x)]. \tag{6.33}$$

Consider a one dimensional wave function and its Fourier transforms

$$\tilde{\Psi}(k) = \mathcal{F}[\Psi(x)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi(x) e^{-ikx} \mathrm{d}x, \tag{6.34}$$

$$\Psi(x) = \mathcal{F}^{-1}[\tilde{\Psi}(k)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\Psi}(k) e^{ikx} \mathrm{d}k, \tag{6.35}$$

yielding for the mean position and wave vector

$$\langle x \rangle = \int_{-\infty}^{\infty} \Psi(x)^* x \Psi(x) \mathrm{d}x, \tag{6.36}$$

$$\langle k \rangle = \int_{-\infty}^{\infty} \tilde{\Psi}(k)^* k \tilde{\Psi}(k) \mathrm{d}k. \tag{6.37}$$

Both are set to zero by a shift in position and frequency space

$$\langle x \rangle = \langle k \rangle = 0, \tag{6.38}$$

resulting for the standard deviations

$$\sigma_x^2 = \langle (x - \langle x \rangle)^2 \rangle = \int_{-\infty}^{\infty} \Psi(x)^* (x - \langle x \rangle)^2 \Psi(x) \mathrm{d}x = \int_{-\infty}^{\infty} \Psi(x)^* x^2 \Psi(x) \mathrm{d}x, \tag{6.39}$$

$$\sigma_k^2 = \langle (k - \langle k \rangle)^2 \rangle = \int_{-\infty}^{\infty} \tilde{\Psi}(k)^* (k - \langle k \rangle)^2 \tilde{\Psi}(k) \mathrm{d}k = \int_{-\infty}^{\infty} \tilde{\Psi}(k)^* k^2 \tilde{\Psi}(k) \mathrm{d}k. \tag{6.40}$$

Multiplying both expressions and using the Plancherel theorem

$$-i \frac{\mathrm{d}}{\mathrm{d}x} \Psi(x) = k \tilde{\Psi}(k), \tag{6.41}$$

we rewrite the first integral as

$$\sigma_x^2 \sigma_k^2 = \int_{-\infty}^{\infty} \Psi(x)^* \left(-i \frac{\mathrm{d}}{\mathrm{d}x} \right)^2 \Psi(x) \mathrm{d}k \int_{-\infty}^{\infty} \Psi(x)^* x^2 \Psi(x) \mathrm{d}x. \tag{6.42}$$

With the Cauchy-Schwarz inequality we have

$$\sigma_x^2\sigma_k^2 \geq \left| \int_{-\infty}^{\infty} -i\frac{\mathrm{d}\Psi(x)}{\mathrm{d}x} x\Psi(x)^* \mathrm{d}x \right|^2 \tag{6.43}$$

$$= \left(\int_{-\infty}^{\infty} i\frac{\mathrm{d}\Psi(x)}{\mathrm{d}x} x\Psi(x)^* \mathrm{d}x \right) \left(\int_{-\infty}^{\infty} -i\frac{\mathrm{d}\Psi(x)^*}{\mathrm{d}x} x\Psi(x) \mathrm{d}x \right) \tag{6.44}$$

$$\geq \mathrm{Im}\left\{ \left(\int_{-\infty}^{\infty} i\frac{\mathrm{d}\Psi(x)}{\mathrm{d}x} x\Psi(x)^* \mathrm{d}x \right) \right\}^2. \tag{6.45}$$

We solve the first integral by partial integration, ($|(\Psi(\pm\infty)|^2 = 0$)

$$\left(\int_{-\infty}^{\infty} i\frac{\mathrm{d}\Psi(x)}{\mathrm{d}x} x\Psi(x)^* \mathrm{d}x \right) = x\Psi(x)\Psi(x)^* \Big|_{-\infty}^{\infty} - i\int_{-\infty}^{\infty} \frac{\mathrm{d}(x\Psi(x)^*)}{\mathrm{d}x} \Psi(x) \mathrm{d}x \tag{6.46}$$

$$= -i\left(\int_{-\infty}^{\infty} \Psi(x)^* \Psi(x) \mathrm{d}x \right) - i\left(\int_{-\infty}^{\infty} x\frac{\mathrm{d}\Psi(x)^*}{\mathrm{d}x} \Psi(x) \mathrm{d}x \right) \tag{6.47}$$

$$= -i + \left(\int_{-\infty}^{\infty} i\frac{\mathrm{d}\Psi(x)}{\mathrm{d}x} x\Psi(x)^* \mathrm{d}x \right)^* \tag{6.48}$$

resulting in

$$-i = -2i\,\mathrm{Im}\left\{ \left(\int_{-\infty}^{\infty} i\frac{\mathrm{d}\Psi(x)}{\mathrm{d}x} x\Psi(x)^* \mathrm{d}x \right) \right\}, \tag{6.49}$$

ergo

$$\mathrm{Im}\left\{ \left(\int_{-\infty}^{\infty} i\frac{\mathrm{d}\Psi(x)}{\mathrm{d}x} x\Psi(x)^* \mathrm{d}x \right) \right\} = \frac{1}{2}. \tag{6.50}$$

Eventually we arrive at the Fourier uncertainty relation

$$\sigma_x^2\sigma_k^2 \geq \mathrm{Im}\left\{ \left(\int_{-\infty}^{\infty} i\frac{\mathrm{d}\Psi(x)}{\mathrm{d}x} x\Psi(x)^* \mathrm{d}x \right) \right\}^2 = \frac{1}{4} \tag{6.51}$$

and with $p = \hbar k$ we arrive at the Heisenberg uncertainty relation [Qui06]

$$\sigma_x^2\sigma_p^2 \geq = \hbar^2\,\mathrm{Im}\left\{ \left(\int_{-\infty}^{\infty} i\frac{\mathrm{d}\Psi(x)}{\mathrm{d}x} x\Psi(x)^* \mathrm{d}x \right) \right\}^2 = \frac{\hbar^2}{4} \tag{6.52}$$

by taking the positive square root

$$\sigma_x\sigma_p \geq \frac{\hbar}{2}. \tag{6.53}$$

At first glance, phase space dynamics in quantum mechanics similar to classical mechanics contradicts the Heisenberg uncertainty relation. It is commonly said that the Heisenberg uncertainty principle is concerned with the measurement process. However its derivation is not concerned with any measuring process, as seen in the preceding.

Historical criticism was raised by Ballentine [BAL70] in 1970. It concerns the nature of the wave function. Ballentine interpreted the wave function not as describing properties of singular systems but as a statistical ensemble of many measurements of equally prepared systems. Thus, it could be possible to break the Heisenberg uncertainty relation for singular experiments. This has not been done experimentally.

Contemporaries of Heisenberg, in particular Herbert Weyl and Eugene Wigner, developed

an approach to make quantum-mechanical calculations in phase space (QMPS). The basis of their formalism is the Wigner function

$$f(x,p) = \frac{1}{2\pi}\int \mathrm{d}y \Psi^*(x - \frac{\hbar}{2}y)e^{-iyp}\Psi(x+\frac{\hbar}{2}y), \tag{6.54}$$

which can be calculated from the wave function of the system or the Moyal equation

$$\frac{\partial f}{\partial t} = \frac{H \star f - f \star H}{i\hbar} := \{\{H, f\}\} \tag{6.55}$$

with the $\star$-product being defined

$$\star = e^{\frac{i\hbar}{2}(\partial_x^{\leftarrow}\partial_p^{\rightarrow} - \partial_p^{\leftarrow}\partial_x^{\rightarrow})}. \tag{6.56}$$

$\{\{\cdot,\cdot\}\}$ is the Moyal Bracket and H is the classical Hamiltonian for the system. Detailed reviews and overviews of QMPS can be found in [CZ12; CFZ16]. $f(x,p)$ is a pseudo-probability distribution, because it can give negative values, in contrast to conventional probability distributions. This counter-intuitiveness and the complicated nature of QMPS may explain its relative obscurity.
With the quantum Hamilton equations (in reduced units),

$$\mathrm{d}x = p\mathrm{d}t + \mathrm{d}W_f \tag{6.57}$$

$$\mathrm{d}p = \partial_x V(x)\mathrm{d}t + \partial_x u(x)\mathrm{d}W_b\ , \tag{6.58}$$

we can propagate X_t forward and P_t separately backward in time. However, it is not trivially possible however to simply reverse the time direction of the backward equation for the momentum. We could solve the position forward in time and then solve the momentum equation separately, but that is not the same as in classical mechanics, where there is no difficulty to propagate both position and momentum forward in time. To find a forward stochastic differential equation for the velocity, which is equivalent to finding one for the momentum due to $p = mu$, we use Itô's lemma

$$\begin{aligned}\mathrm{d}u(x(t)) &= \partial_x u(x(t))\mathrm{d}x(t) + \frac{1}{2}\partial_x^2 u(x(t))\mathrm{d}t \\ &= \left[\partial_x u(x(t))u(x(t)) + \frac{1}{2}\partial_x^2 u(x(t))\right]\mathrm{d}t + \partial_x u(x(t))\mathrm{d}W_f,\end{aligned} \tag{6.59}$$

and with replacing the stochastic process $u(x(t))$ with the stochastic process of the momentum itself U_t, we arrive at the forward stochastic differential equation for the momentum

$$\mathrm{d}U_t = \left[\partial_x u(x(t))U_t + \frac{1}{2}\partial_x^2 u(x(t))\right]\mathrm{d}t + \partial_x u(x(t))\mathrm{d}W_f. \tag{6.60}$$

For the harmonic oscillator in reduced units with $u(x) = -x$ the position and momentum equations decouple:

$$\mathrm{d}U_t = -U_t\mathrm{d}t - \mathrm{d}W_f. \tag{6.61}$$

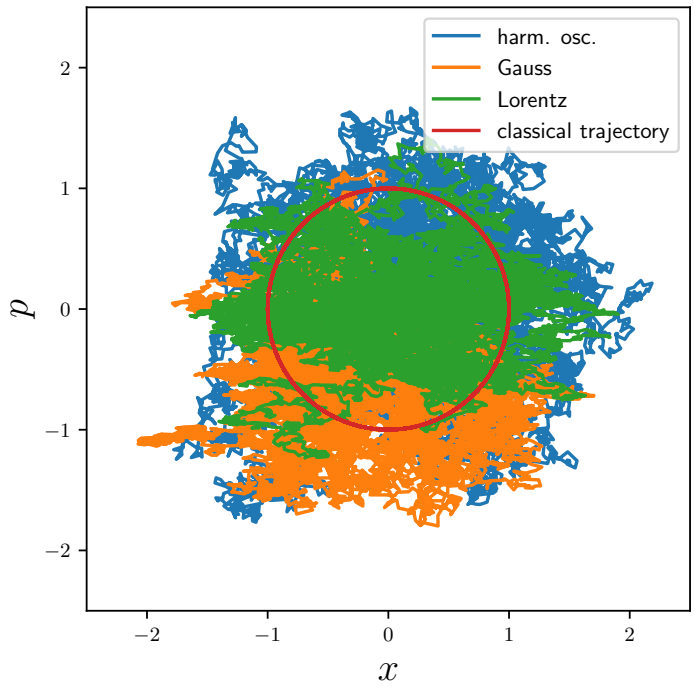


FIGURE 6.7: Phase space sample paths for the three potentials. For the ground state, we have an Ornstein-Uhlenbeck process for the position and momentum(see eqs. (6.62) and (6.61)) for harmonic oscillator and similar processes for the Gauss and Lorentz potential (see eqs. (6.63) and (6.64)). The equivalent classical trajectory of the harmonic oscillator, which is a circle in reduced units is also drawn to highlight differences between classical and quantum mechanical trajectories.

This is the Ornstein-Uhlenbeck process, the only stationary Gaussian Markov process. We obtain the same process for the position of the harmonic oscillator,

$$\mathrm{d}X_t = -X_t \mathrm{d}t + \mathrm{d}W_f. \tag{6.62}$$

The sign in front of the diffusion term is not important because of the symmetry of the Gaussian distribution $\mathcal{N}(0, \mathrm{d}t)$. Their respective position and momentum distributions are identical Gaussians, consistent with the minimum uncertainty of the ground state of the quantum harmonic oscillator.

The velocity and position processes of the Gauss and Lorentz potential do not decouple:

$$\mathrm{d}X_t = u(X_t)\mathrm{d}t + \mathrm{d}W_f, \tag{6.63}$$

$$\mathrm{d}U_t = \left(\frac{\partial u(X_t)}{\partial x} U_t + \frac{1}{2}\frac{\partial^2 u(X_t)}{\partial x^2}\right)\mathrm{d}t + \frac{\partial u(X_t)}{\partial x}\mathrm{d}W_f. \tag{6.64}$$

There, uppercase letters indicate the value of the corresponding stochastic process at time t and lowercase letters indicate the functions, obtained beforehand, of the stochastic position process. It has to be mentioned, that these paths are sample paths, possible paths of the particles.

These sample paths can be see in Fig. 6.7 with the corresponding probability densities in Fig. 6.8. The probability densities are consistent with what would be expected for the three

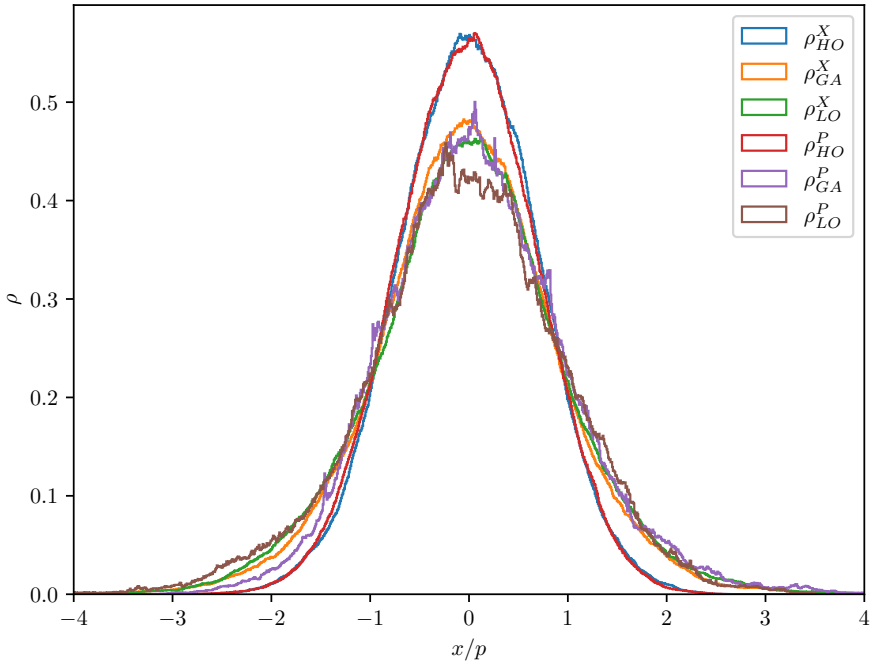


FIGURE 6.8: Probability distributions for position and momentum of each of the three potentials. For the harmonic oscillator, they overlap quite well as expected in reduced units. The distribution of the position for the Gaussian and Lorentzian potentials and for their momenta also overlap but are broader.

potentials with the densities of the harmonic oscillator almost completely overlapping, only deviating from one another slightly due to numerical errors. The probability density of the position of the other two potentials overlap, as do the distributions of the momentum. Both distributions of the other two potentials are broader than that of the harmonic oscillator, resulting in a larger uncertainty (see Tab. 6.2). It has to be mentioned that the ground states of quantum mechanical (harmonic) oscillators do not oscillate. This is consistent with conventional quantum mechanics. If one wants to observe quantum mechanical oscillators that acually oscillate, one has to look at coherent states.

Potential	$\Delta x \Delta p$	$\left\lvert \frac{\Delta x \Delta p}{0.5} - 1 \right\rvert$
harm. Osc.	0.48500	0.03
Gauss	0.83675	0.6735
Lorentz	0.93271	0.86542

TABLE 6.2: Heisenberg uncertainties of the ground states of the three potentials as multiples of the reduced Planck constant $\hbar$ and the difference to the expected harmonic oscillator value, the minimal uncertainty ($\frac{\hbar}{2}$). One can see that the Gauss and Lorentz potential yield larger uncertainty than the harmonic oscillator. The uncertainty of the harmonic oscillator being a little bit less than the minimal uncertainty is a numerical error.

6.7 Coherent States

Coherent states, also called Glauber states, first proposed by Schrödinger in 1926 [Sch26a], are bound, non-stationary states of the harmonic oscillator, that in the classical limit behave like the equivalent classical state of the harmonic oscillator [Gla63; KS85]. Coherent states can be written either as a series of superposing eigenstates with [Per77]

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{6.65}$$

or with a unitary operator acting on the ground state [Per77]

$$e^{\alpha \hat{a}^\dagger - \alpha^* \hat{a}} |0\rangle . \tag{6.66}$$

Coherent states are eigenstates of the lowering operator

$$\hat{a} |\alpha\rangle = \alpha |\alpha\rangle . \tag{6.67}$$

In [PB13] it is shown that from the classical phase space trajectories $x_{cl}(t), p_{cl}(t)$ one can derive the probability density

$$\rho(x,t) = (2\pi\tilde{\sigma})^{\frac{1}{2}} e^{-\frac{(x-x_{cl}(t))^2}{2\tilde{\sigma}}} \tag{6.68}$$

with $\tilde{\sigma} = \frac{\hbar}{2m\omega}$ and the action

$$S(x,t) = x p_{cl}(t) - \frac{1}{2} x_{cl}(t) - \frac{1}{2}\hbar\omega t. \tag{6.69}$$

The drift and osmotic velocities

$$v(x,t) = \frac{p_{cl}(t)}{m}, \tag{6.70}$$

$$u(x,t) = -\omega(x - x_{cl}(t)), \tag{6.71}$$

yield for the propagation in space

$$\mathrm{d}x = \left(\frac{p_{cl}(t)}{m} - \omega(x - x_{cl}(t)) \right) \mathrm{d}t + \sigma \mathrm{d}W_f. \tag{6.72}$$

Employing Itô's lemma and using

$$p = m(v(X_t,t) + u(X_t,t)) \tag{6.73}$$

we obtain

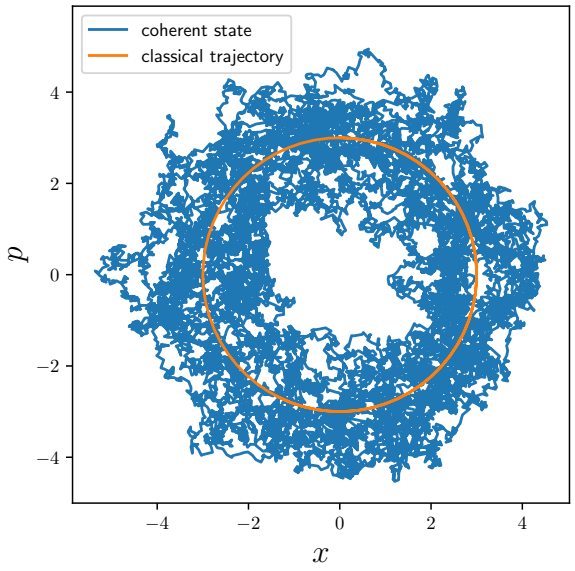
$$\begin{aligned} \mathrm{d}P_t = & \left(\frac{\partial p}{\partial x} \left(\frac{p_{cl}(t)}{m} - \omega(x - x_{cl}(t)) \right) \right. \\ & \left. + M \left(\frac{\partial x_{cl}(t)}{\partial t} + \frac{1}{M} \frac{\partial p_{cl}(t)}{\partial t} \right) \right) \mathrm{d}t + \frac{\partial p}{\partial x} \mathrm{d}W_f \qquad (6.74) \\ = & \left(-\omega P_t + M \left(\frac{\partial x_{cl}(t)}{\partial t} + \frac{1}{M} \frac{\partial p_{cl}(t)}{\partial t} \right) \right) \mathrm{d}t - \omega \mathrm{d}W_f. \qquad (6.75) \end{aligned}$$

The starting points used were $X_0 = 0, P_0 = 3$ which yield the classical trajectories

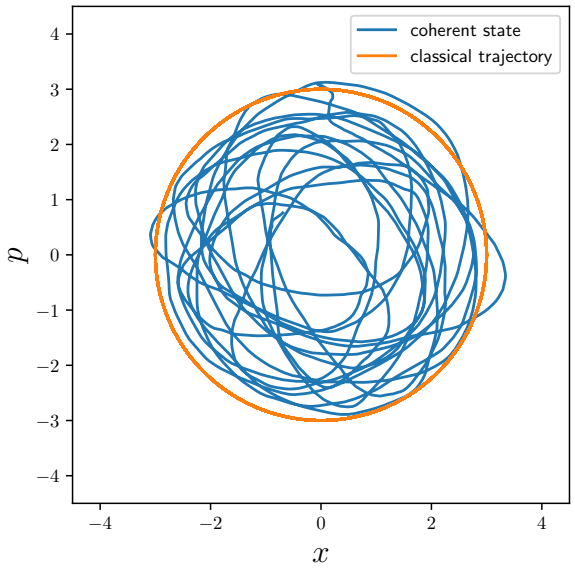
$$x_{cl}(t) = x_0 \cos \omega t + \frac{p_0}{M\omega} \sin \omega t = 3 \sin(\omega t), \tag{6.76}$$

$$p_{cl}(t) = p_0 \cos \omega t - M\omega x_0 \sin \omega t = -3 \cos(\omega t). \tag{6.77}$$

This results in a coherent state being a quantum mechanical ground state with a superposed classical motion in phase space. The resulting trajectory is shown in Fig. 6.9a.



(A)



(B)

FIGURE 6.9: Phase space trajectory with(bottom) and without (top) time-moving average for smoothing applied. The phase portrait with the time-moving average looks quite similar to a classical oscillator.

In Fig. 6.9b one can see, that the result is much closer to a classical oscillator due to an applied time-moving average. The probability densities match the expected values very closely.
Coherent states are minimal uncertainty states. This can be shown with the ladder operators for the position and momentum operators

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a} + \hat{a}^\dagger), \tag{6.78}$$

$$\hat{p} = -i\sqrt{\frac{m\omega\hbar}{2}}(\hat{a} - \hat{a}^\dagger), \tag{6.79}$$

$$\tag{6.80}$$

thus resulting in the mean position and momentum

$$\langle x\rangle = \sqrt{\frac{\hbar}{2m\omega}}\,\langle\alpha|\,(\hat{a} + \hat{a}^\dagger)\,|\alpha\rangle = \sqrt{\frac{\hbar}{2m\omega}}(\alpha + \alpha^*), \tag{6.81}$$

$$\langle p\rangle = -i\sqrt{\frac{m\omega\hbar}{2}}\,\langle\alpha|\,(\hat{a} - \hat{a}^\dagger)\,|\alpha\rangle = -i\sqrt{\frac{m\omega\hbar}{2}}(\alpha - \alpha^*), \tag{6.82}$$

as well as the second moments

$$\langle x^2\rangle = \frac{\hbar}{2m\omega}((\alpha + \alpha^*)^2 + 1), \tag{6.83}$$

$$\langle p^2\rangle = -\frac{m\omega\hbar}{2}((\alpha - \alpha^*)^2 - 1) \tag{6.84}$$

with the uncertainty

$$\Delta x\Delta p = \sqrt{\langle x^2\rangle - \langle x\rangle}\sqrt{\langle p^2\rangle - \langle p\rangle} = \frac{\hbar}{2}, \tag{6.85}$$

if t is known in regard to the classical motion in phase space. This finding is consistent with the coherent state being a quantum mechanical ground state superposed with classical motion in phase space, with the latter naturally having an uncertainty of zero. Integrating and normalizing equation (6.68) over time, we get a uncertainty larger than the minimal one (see Fig. 6.10).

To verify the correctness of this stochastic ansatz concerning coherent states, one can increase the energy of the coherent state by changing the starting point and thus the classical phase space trajectory. Consistent with the correspondence principle, the larger the energy, the more the classical behavior should dominate over the quantum mechanical ground state, with the latter ones energy being naturally fixed at $E_0 = \frac{\hbar\omega}{2}$. This domination of the classical behavior over the quantum mechanical one should be visible in the phase portrait. Changing the starting points to be $X_0 = 0, P_0 = -100$, we increase the energy of the classical trajectory by two orders of magnitude. The resulting coherent phase space trajectory is almost identical to the corresponding classical phase space trajectory (see Fig. 6.11).

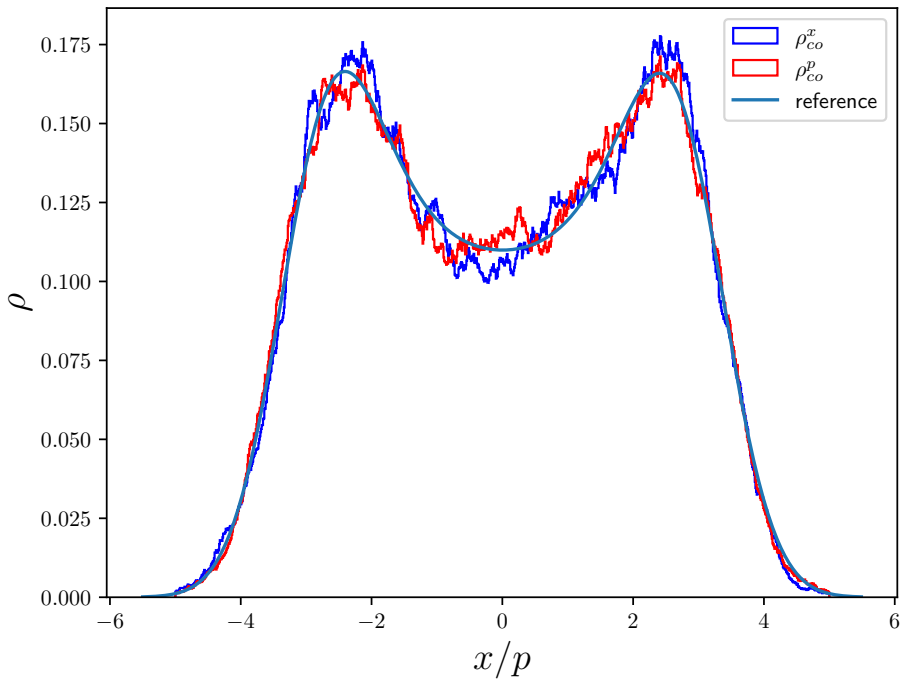


FIGURE 6.10: Probability densities for the position and momentum correspond with the expected values. Even though coherent states are minimal uncertainty states [Gla63], if the time t is that of the classical motion, here we integrate over time, finding larger uncertainty than $\frac{\hbar}{2}$.

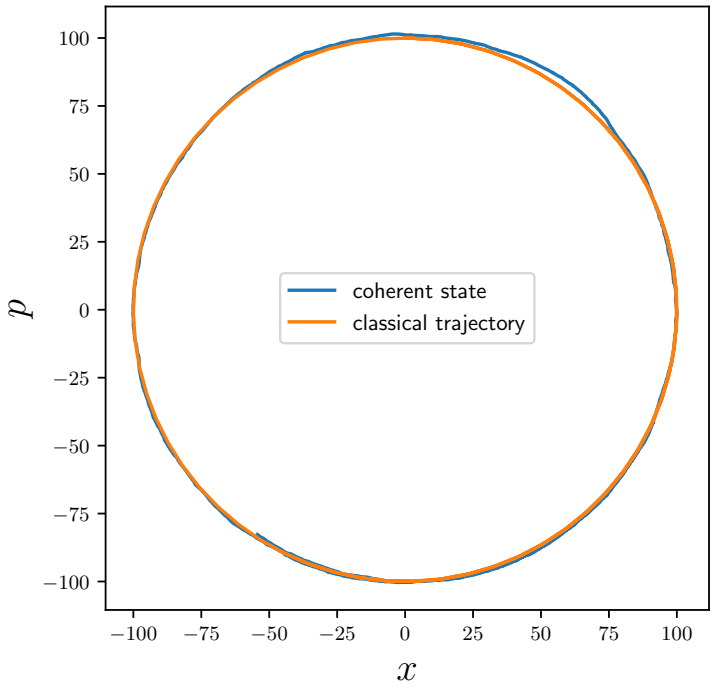


FIGURE 6.11: Coherent and classical phase space trajectory for much larger energies than in the initial simulation. The energy is increased by changing the starting points to $X_0 = 0, P_0 = -100$ in reduced units. For higher energies, the classical behavior dominates over the ground state fluctuations, which is to be expected.

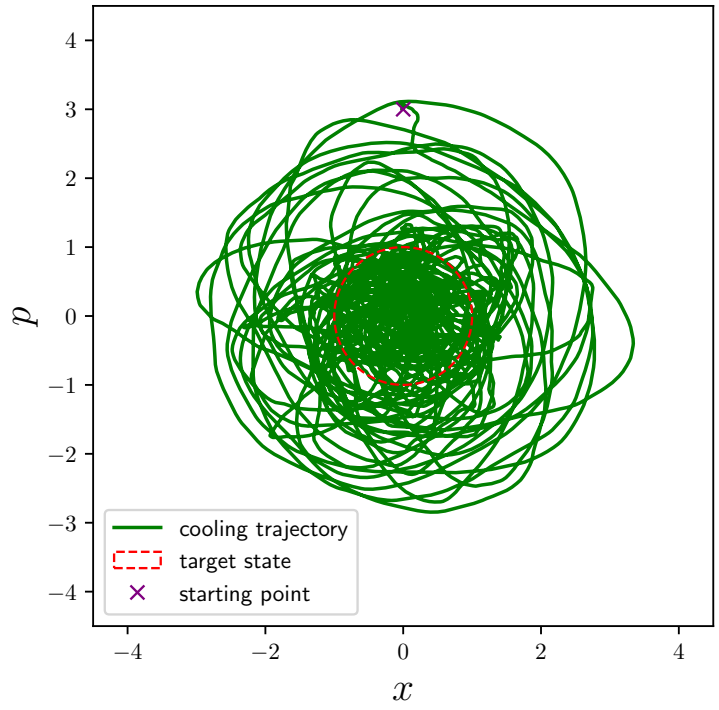


FIGURE 6.12: Phase space trajectory of the cooling process with the time-moving average applied.

6.7.1 Modeling the Cooling Process

In [Mag+21], the phase space trajectory along the cooling process, where the energy of the system is decreased to get near the ground state, is also shown. To model this, we replaced the conservative, classical trajectory of the coherent state by that of a damped harmonic oscillator

$$x_{cool}(t) = x_{cl}(t) \cdot e^{-t/t_d} = 3\sin(\omega t) \cdot e^{-t/t_d} \tag{6.86}$$

$$p_{cool}(t) = p_{cl}(t) \cdot e^{-t/t_d} = 3\cos(\omega t) \cdot e^{-t/t_d}, \tag{6.87}$$

with the damping time equal to the reaction time of Kalman filter, $t_d = 32$ ns. The shift of the oscillation frequency can be neglected for this choice of damping. The result can be seen in Fig. 6.12. The cooling trajectory closely resembles the cooling trajectory for energy $E = 2.71\hbar\omega$ shown in [Mag+21], although here we cool down completely to the quantum mechanical ground state.

Chapter 7

Conclusion

The stochastic mechanics description of quantum mechanics stores all information about the system at hand in two velocity fields, the current velocity $v(x,t)$ and the osmotic velocity $u(x,t)$, like the wave function $\Psi(x,t)$ in conventional quantum mechanics. The wave function solves the Schrödinger equation which is comparatively easy to solve, being a partial differential equation and an eigenvalue problem in the stationary case. Calculating the velocity fields is however a non-trivial task. If one does not want to rely on the Schrödinger equation, one had to find the Nash equilibrium of the quantum Hamilton equations by using the Bender/Steiner algorithm.

In this thesis, I have shown that the Pavon principle is equivalent to the Rayleigh-Ritz principle of conventional quantum mechanics, if one considers the stationary ground state of a quantum system. This fact was applied in a genetic algorithm which uses neural networks to parametrize the osmotic velocity. to minimize first the energy functional and then solving the Ricatti equation. Like in conventional quantum mechanics, solving the ground state of one-dimensional systems is enough to obtain all exited states with corresponding eigenenergies by finding the ground states of the supersymmetric partner Hamiltonians. This ansatz of finding the osmotic velocity has a lot of advantages compared to the Bender/Steiner algorithm: It is more efficient, easier to program, yields better results along a larger part of the coordinate space, and is conceptually easier to understand. This algorithm was applied to solve the harmonic oscillator, the Pöschl-Teller potential and the quartic double-well as a proof of concept. The ground states and the exited states were solved together with the eigenenergies.

A big advantage of the stochastic description of quantum mechanics is its relation to the concept of time. Conventional quantum mechanics lacks a time-measurement operator. Many workarounds were proposed over the years, but none yielded results that matched experimental results. Having stochastic trajectories X_t makes time measurements a simple task, by solving the stochastic differential equation and analyzing the results.

The proposed algorithm was used to obtain the osmotic velocities of three potentials, the Gauss potential, the Lorentz potential and the harmonic oscillator. With the osmotic velocity, the stochastic differential equations were solved. Then using methods from time-series analysis, the autocorrelation function and the power spectral density, the results were compared to each other and to experimental results. In the experiment, optical tweezers were used to levitate a spherical nanoparticle. The two potentials acting on that particle were the Gauss and the Lorentz potential, which were both approximated by a harmonic oscillator.

The simulations yielded no experimentally measurable, significant differences in the autocorrelation functions and the power spectral densities, but larger differences in the ground state energy among the three potential forms.

The Heisenberg uncertainties $\Delta x \Delta p$ were also calculated. Because time reversing the momentum equation of the quantum Hamilton equation is a non-trivial task, the Itô formula was used to derive stochastic differential equations for the osmotic velocity process, which is the momentum if multiplied with the mass of the particle in the ground state, forward in time. This yielded two decoupled Ornstein-Uhlenbeck processes for the harmonic oscillator and two coupled stochastic differential equations for the Gauss and Lorentz potential. These were integrated forward in time, resulting in sample phase space trajectories, which however still obey the Heisenberg uncertainty relation. The calculated Heisenberg uncertainties for the Gauss and Lorentz potential were much larger than the one of the harmonic oscillator. Due to the experimental results for higher energy states in the phase portraits of the sample phase space trajectories, this ansatz was applied to coherent states that are non-stationary, bounded states of the harmonic oscillator with classical behavior in the classical limit and for modeling the cooling process therefore confirming the correspondence principle.

7.1 Outlook

Being one of the youngest descriptions of quantum mechanics, stochastic mechanics has a lot of open problems, that are already solved in conventional quantum mechanics.
Even though a theoretical extension to multidimensional problems is possible, there is still a lack of methods that solve such problems effectively, in contrast to conventional quantum mechanics (for example density functional theory). The algorithm proposed in this thesis, could be extended to multidimensional systems. It has yet to be determined, how much the curse of dimensionality influences the performance of the algorithm.
A description of relativistic problems is also yet to be accomplished. Stochastic mechanics on curved manifolds was used to describe the spin of particles[BP23; Bey23]. Thus, a description on Lorentzian manifolds should theoretically be possible.
Treating a system of two particles with spin $\frac{1}{2}$ should also be an endeavor worth pursuing, especially considering that there might be stable bound states between such particles like in the hydrogen atom, where the influence of the spins is usually disregarded.
Methods from time-series analysis can be used to support future experiments but also have the potential of gaining new insight into quantum systems, as do phase space dynamics. The connection between QMPS and the phase space description derived in this thesis should also be expanded on.

The results of this thesis extend the foundation of Nelson's description of quantum mechanics and its applications and possibilities. There is still a lot of work to be done in this field of research.

Appendix A

Elements of Stochastics

A.1 The p-Limit

The p-limit is defined for a sequence $X_n \in L^p$ with $X \in L^p$ being a convergence value. If

$$\mathbb{E}[|X_n - X|^p] \xrightarrow{n\to\infty} 0 \tag{A.1}$$

then the sequence X_n converges in the p-th mean towords X. If $p = 2$ then one can also write

$$ms\text{-}lim_{n\to\infty} X_n = X \tag{A.2}$$

meaning the sequence converges in the mean square towards X. If $p = 1$, the sequence converges in the mean [Kö18].

A.2 $\mathcal{F}$-measurable Functions

Consider a probability space $(\Omega, \mathcal{F}, \mathcal{P})$, with Ω denoting the set of all possible outcomes, $\mathcal{F}$ being the σ-algebra of subsets of Ω, and $\mathcal{P}$ assigning a probability from 0 to 1 to all events. A function $f : \Omega \to \mathbb{R}^d$ is $\mathcal{F}$-measurable, if for the inverse function and U an open subset of $\mathbb{R}^d$

$$f^{-1}(U) := \{\omega \in \Omega; f(\omega) \in U)\} \in \mathcal{F} \tag{A.3}$$

holds [Kö18]. This means that the preimage (or inverse image) of f lies within the event space.

A.3 Stratonovich Integrals

Next to the Itô integrals, there exists another class of stochastic integrals, the Stratonovich integrals [Øks10]

$$\int_0^t Y(x(s), s) \circ \mathrm{d}W_s := ms\text{-}lim_{n\to\infty} \sum_{i=1}^{n} Y\left(x\left(\frac{t_{i-1} + t_i}{2}\right), \frac{t_{i-1} + t_i}{2}\right) \tag{A.4}$$

denoted with $\circ$ to differentiate it from Itô integrals, which can be used to write a Stratonovich stochastic differential equation

$$x(t) = x(t_0) + \int_{t_0}^{t} \alpha(x(s), s)\mathrm{d}s + \int_{t_0}^{t} \beta(x(s), s) \circ \mathrm{d}W_s. \tag{A.5}$$

One can transform Stratonovich integrals into Itô integrals with

$$\int_{t_0}^{t} \beta(x(s),s) \circ \mathrm{d}W_s = \int_{t_0}^{t} \beta(x(s),s)\mathrm{d}W_s + \frac{1}{2}\int_{t_0}^{t} b(x(s),s)\partial_x \beta(x(s),s)\mathrm{d}s, \tag{A.6}$$

in which $x(s)$ are the solutions of the connected Itô stochastic differential equation. Using this and setting

$$\alpha(x(s),s) = a(x(s),s) - \frac{1}{2}b(x(s),s)\partial_x b(x(s),s), \tag{A.7}$$

$$\beta(x(s),s) = b(x(s),s), \tag{A.8}$$

one can see that the Itô stochastic differential equation

$$\mathrm{d}x = a\mathrm{d}t + b\mathrm{d}W_t \tag{A.9}$$

is equivalent to

$$\mathrm{d}x = \left(a - \frac{1}{2}b\partial_x b\right)\mathrm{d}t + b \circ \mathrm{d}W_t, \tag{A.10}$$

the corresponding Stratonovich stochastic differential equation [Øks10]. Even though both integrals are equivalent, only Itô integrals and stochastic differential equations are used in this thesis.

A.4 Convergence Orders

For numerical integrators of stochastic differential equations, one usually distinguishes between weak and strong convergence. The latter are defined with regard to two constants K_T and $K_{g,T}$, the step size Δ and arbitrary polynomials $g(\cdot)$. The strong convergence order is defined as

$$\mathbb{E}(|X(T) - X_{N_T}|) \leq K_T \Delta^{\gamma} \tag{A.11}$$

and the weak order is defined as

$$|\mathbb{E}(g(X_T)) - \mathbb{E}(g(X_T))| \leq K_{g,T}\Delta^{\beta}, \tag{A.12}$$

with the largest possible values of γ and β giving the corresponding strong and weak orders [Klo02; Sch10].
Strong order depends on the mean of the difference between the estimated and the exact solution and the weak order depends on the difference of the mean of the estimated and the mean of the exact solution. For comparison with deterministic integration schemes, one would use the strong convergence order. Stochastic integration methods usually have a low convergence order compared with their deterministic counterparts [Haa21]. However, in physics it is more common to write down the error of a numerical scheme with regard to the growth of the error, denoted as $\mathcal{O}(\Delta^n)$. For the solution methods for stochastic differential equations, each indicator γ and β can be seen in table A.1.

scheme	γ	β	$\mathcal{O}$
Euler-Maruyama	$\frac{1}{2}$	1	$\mathcal{O}(\sqrt{\Delta})$ [Pat18]
Heun	1	1	$\mathcal{O}(\Delta^{\frac{3}{2}})$ [Bog+23]
Milstein	1	1	$\mathcal{O}(\Delta^2)$ [Mil75]

TABLE A.1: Convergence orders [Klo02] and errors for the three most common numerical integration schemes for Itô-stochastic differential equations.

A.5 Milstein Sheme

The Milstein sheme is another numerical method of solving stochastic differential equations [Sch99]. It adds a term to the Euler-Scheme resulting in

$$X_{n+1} = X_n + a\Delta t + b\Delta W_n + \frac{1}{2}bb'\Big((\Delta W)^2 - \Delta t\Big). \tag{A.13}$$

It is generally better than the Euler-Maruyama scheme with a strong convergence order of 1 [Klo02] compared to Euler-Maruyamas of $\frac{1}{2}$ [Sch99]. However, if the diffusion coefficient b is a constant, as is the case for the quantum Hamilton equation for position, the additional term becomes zero and thus doesn't improve on the Euler scheme. That is, why for this thesis the Heun scheme was chosen (also convergence orders of 1 [Kö18]).

A.6 Derivation of the Fokker-Planck Equation from the Itô Formula

Itô diffusion processes

$$\mathrm{d}X_t = a(X,t)\mathrm{d}t + b(X,t)\mathrm{d}W_t \tag{A.14}$$

have a variety of valid solutions $(X_{t_0}, \dots X_T)$. Alternatively, one can solve such a stochastic differential equation by considering the probability distribution $\rho(x,t)$, giving the probability of being at time t at x. To find $\rho(x,t)$ one has to solve the Fokker-Planck equation

$$\partial_t\rho(x,t) = -\partial_{x_i}(a(x,t)\rho(x,t)) + \frac{1}{2}\partial_{x_i}\partial_{x_j}(b_{i,k}(x,t)b_{j,k}(x,t)\rho(x,t)) \tag{A.15}$$

or for one dimension

$$\partial_t\rho(x,t) = -\partial_x(a(x,t)\rho(x,t)) + \frac{1}{2}\partial_x^2(b^2(x,t)\rho(x,t)). \tag{A.16}$$

This is a partial differential equation, which can be derived using the Itô formula [Bey23; Gar85].
Considering a smooth function $f(X_t)$ of X_t undergoing an Itô diffusion process, we can use the Itô formula

$$\mathrm{d}f = \partial_x f\mathrm{d}X_t + \frac{1}{2}\partial_x^2 f\mathrm{d}X_t^2 \tag{A.17}$$

$$= \partial_x f(a(X,t)\mathrm{d}t + b(X,t)\mathrm{d}W_t) + \frac{1}{2}b^2\mathrm{d}t \tag{A.18}$$

and after separating the drift and diffusion term we get

$$\mathrm{d}f = (\partial_x f a(X,t) + \frac{1}{2}b^2\partial_x^2)\mathrm{d}t + \partial_x f b(X,t)\mathrm{d}W_t. \tag{A.19}$$

Now taking the expectation of the equation and using $\mathbb{E}[\mathrm{d}W_t] = 0$, we obtain

$$\mathbb{E}[\mathrm{d}f] = \mathbb{E}[\partial_x f a(X,t) + \frac{1}{2}b^2\partial_x^2 f]\mathrm{d}t, \tag{A.20}$$

and by dividing by $\mathrm{d}t$, we arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[f] = \mathbb{E}[\partial_x f a(X,t) + \frac{1}{2}b^2\partial_x^2 f]. \tag{A.21}$$

Now, introducing the probability distribution $\rho(x,t)$. We rewrite the expectation values with integrals over all possible positions,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} f(x)\rho(x,t)\mathrm{d}x = \int_{-\infty}^{\infty}(a(x,t)\partial_x f(x) + \frac{b^2(x,t)}{2}\partial_x^2 f(x))\rho(x,t)\mathrm{d}x \tag{A.22}$$

$$= \int_{-\infty}^{\infty} a(x,t)\partial_x f(x)\rho(x,t)\mathrm{d}x + \int_{-\infty}^{\infty}\frac{b^2(x,t)}{2}\partial_x^2 f(x)\rho(x,t). \tag{A.23}$$

Both terms can now be treated with partial integration $\int_a^b uv'\mathrm{d}x = uv|_a^b - \int_a^b u'v\mathrm{d}x$ to get rid of the partial derivatives of $f(x)$ in both terms on the right side of the equation. For the first term, we get

$$\int_{-\infty}^{\infty} a(x,t)\partial_x f(x)\rho(x,t)\mathrm{d}x = f(x)a(x,t)\rho(x,t)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(x)\partial_x(a(x,t)\rho(x,t))\mathrm{d}x \tag{A.24}$$

with

$$f(x)a(x,t)\rho(x,t)\Big|_{-\infty}^{\infty} = 0. \tag{A.25}$$

We have to partially integrate the second term twice. Again, only the integral will remain and the rest becomes zero, thus the second term becomes

$$\int_{-\infty}^{\infty} b^2(x,t)\partial_x^2 f(x)\rho(x,t)\mathrm{d}x = \int_{-\infty}^{\infty} f(x)\partial_x^2(b^2(x,t)\rho(x,t))\mathrm{d}x. \tag{A.26}$$

Now we replace the terms in equation (A.23) and get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} f(x)\rho(x,t)\mathrm{d}x = &-\int_{-\infty}^{\infty} f(x)\partial_x(a(x,t)\rho(x,t))\mathrm{d}x \\ &+\frac{1}{2}\int_{-\infty}^{\infty} f(x)\partial_x^2(b^2(x,t)\rho(x,t))\mathrm{d}x.\end{aligned} \tag{A.27}$$

Rearranging the time derivative of the right side into the integral, it will only act on the probability density. Then we put all the terms on one side and combine the integrals to obtain

$$\int_{-\infty}^{\infty} f(x)\left(\frac{\partial}{\partial t}\rho(x,t) + \partial_x(a(x,t)\rho(x,t)) - \frac{1}{2}\partial_x^2(b^2(x,t)\rho(x,t))\right)\mathrm{d}x = 0. \tag{A.28}$$

Because $f(x)$ is arbitrary, the bracketed part of the integral has to be zero, thus

$$\frac{\partial}{\partial t}\rho(x,t) = -\partial_x\big(a(x,t)\rho(x,t)\big) + \frac{1}{2}\partial_x^2\big(b^2(x,t)\rho(x,t)\big), \tag{A.29}$$

which is the Fokker-Planck equation for an arbitrary Itô diffusion process.
The Fokker-Planck equation is connected to the continuity equation. By taking out one derivative in space from the right hand side, one gets

$$\frac{\partial}{\partial t}\rho(x,t) = -\partial_x\left[(a(x,t)\rho(x,t)) - \frac{1}{2}\partial_x(b^2(x,t)\rho(x,t))\right]. \tag{A.30}$$

Defining the terms in the bracket as the probability flow, one gets

$$\frac{\partial}{\partial t}\rho(x,t) = -\frac{\partial}{\partial x}j, \tag{A.31}$$

the continuity equation. Alternatively, one could derive the Fokker-Planck equation from the Langevin equation [RF12] or the Chapman-Kolomorov equation [SP10].

Appendix B

Classical Mechanics

B.1 Derivation of the Euler-Lagrange equation from Variational Calculus

The Hamilton principle is

$$\delta S = \delta \int_{t_0}^{t_1} \mathcal{L}(t,q,\dot{q})\mathrm{d}t = 0. \tag{B.1}$$

To obtain the Euler-Lagrange equation, we use variational calculus. For simplicity, we restrict the derivation to one-dimensional problems. We start wit

$$0 = \int_{t_0}^{t_1} \delta\mathcal{L}(t,q,\dot{q})\mathrm{d}t. \tag{B.2}$$

Variation in this context means that a path $q(t)$ is changed by a small amount

$$\tilde{q}(t) = q(t) + \varepsilon\eta(t) \tag{B.3}$$

with

$$\eta(t_0) = \eta(t_1) = 0. \tag{B.4}$$

Inserting this into the Lagrangian

$$\mathcal{L}(t, \tilde{q}, \dot{\tilde{q}}) = \mathcal{L}\left(t, q(t) + \varepsilon\eta(t), \dot{q}(t) + \varepsilon\frac{\mathrm{d}\eta(t)}{\mathrm{d}t}\right), \tag{B.5}$$

the variation becomes a derivative with respect to ε

$$\delta S = \int_{t_0}^{t_1} \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{L}\left(t, q(t) + \varepsilon\eta(t), \dot{q}(t) + \varepsilon\frac{\mathrm{d}\eta(t)}{\mathrm{d}t}\right)\mathrm{d}t \tag{B.6}$$

which becomes using the chain rule

$$\delta S = \int_{t_0}^{t_1} \frac{\partial L}{\partial q}\frac{\partial q}{\partial \varepsilon} + \frac{\partial \mathcal{L}}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial \epsilon}\mathrm{d}t. \tag{B.7}$$

Using partial integration for the second term gives

$$\delta S = \int_{t_0}^{t_1} \frac{\partial L}{\partial q}\frac{\partial q}{\partial \varepsilon} + \frac{\partial \mathcal{L}}{\partial \dot{q}}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial q}{\partial \varepsilon}\mathrm{d}t \tag{B.8}$$

$$= \int_{t_0}^{t_1} \frac{\partial L}{\partial q}\frac{\partial q}{\partial \varepsilon} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}}\frac{\partial q}{\partial \varepsilon}\mathrm{d}t + \frac{\partial L}{\partial \dot{q}}\frac{\partial q}{\partial \varepsilon}\Big|_{t_0}^{t_1} \tag{B.9}$$

where the last term becomes zero due to (B.4). We treat the remaining integral

$$\delta S = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}}\right)\frac{\partial q}{\partial \varepsilon}\mathrm{d}t = 0 \tag{B.10}$$

which can only be valid, if

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \tag{B.11}$$

which is the Euler-Lagrange equation. For more than one dimension, one simply has an Euler-Lagrange equation for each generalized coordinate q_n

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q_n}} - \frac{\partial L}{\partial q_n}. \tag{B.12}$$

B.2 Inequalities of the Classical Mechanics Formulations

It is commonly believed that all formalisms of classical mechanics are equivalent. Even though this statement is generally true, there exist edge cases, where this statement is false. For starters, not all Newtonian systems can be described with Lagrangian or Hamiltonian mechanics, because the force is a vector function $\vec{F} : \mathbb{R}^n \to \mathbb{R}^n$ and the Hamiltonian and the Lagrangian are scalar functions $\mathcal{H}/\mathcal{L} : \mathbb{R}^n \to \mathbb{R}$, ergo, there doesn't exist a homeomorphic map from one to another [CA21].
However all Lagrangian systems are Newtonian. This can be shown, by considering the

Euler-Lagrange equation

$$\frac{\partial L}{\partial x_i} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial v_i} \tag{B.13}$$

$$= \frac{\mathrm{d}x_j}{\mathrm{d}t}\frac{\partial^2 \mathcal{L}}{\partial x_j \partial v_i} + \frac{\mathrm{d}v_k}{\mathrm{d}t}\frac{\partial^2 \mathcal{L}}{\partial v_k \partial v_i} \tag{B.14}$$

$$= v_j \frac{\partial^2 \mathcal{L}}{\partial x_j \partial v_i} + a_k \frac{\partial^2 \mathcal{L}}{\partial v_k \partial v_i}, \tag{B.15}$$

with the last term being equal to

$$a_k \frac{\partial^2 \mathcal{L}}{\partial v_k \partial v_i} = \frac{\partial L}{\partial x_i} - \frac{\partial^2 \mathcal{L}}{\partial x_j \partial v_i} v_j. \tag{B.16}$$

Thus we can obtain the acceleration a_k if the Hessian of the Lagrangian is invertible, which is necessary for the stationary action having a unique solution [CA21].
There are however Hamiltonian systems, that can not be described using Newtonian mechanics. For showing this, we derive a formula for the acceleration with the canonical equations

$$a_i = \frac{\mathrm{d}v_i}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{H}}{\partial p_i} = \frac{\partial^2 \mathcal{H}}{\partial q_j \partial p_i}\frac{\mathrm{d}q_j}{\mathrm{d}t} + \frac{\partial^2 \mathcal{H}}{\partial p_k \partial p_i}\frac{\mathrm{d}p_k}{\mathrm{d}t} \tag{B.17}$$

$$= \frac{\partial^2 \mathcal{H}}{\partial q_j \partial p_i}\frac{\partial \mathcal{H}}{\partial p_j} - \frac{\partial^2 \mathcal{H}}{\partial p_k \partial p_i}\frac{\partial \mathcal{H}}{\partial q_k} \tag{B.18}$$

ergo we obtain a formulation of the acceleration in terms of position and momentum, not velocity. To get that expression, the Hessian of the Hamiltonian to be invertible

$$|\partial_{p_i} v_j| = |\partial_{p_i}\partial_{p_j}\mathcal{H}| \neq 0, \tag{B.19}$$

which is not required from the Hamilton formalism. Ergo, we need an additional property for systems, that can be described in both formalism, the kinematic equivalence, the recoverability of the dynamics from the kinematics of a system [CA21].
Hamiltonian and Lagrangian systems can be easily transformed into one another by the Legendre transformation. This is only possible, if

$$|\partial_{v_i} p_j| = |\partial_{v_i}\partial_{v_j}\mathcal{L}| \neq 0, \tag{B.20}$$

the Hessian of the Lagrangian is invertible. Their determinants are also connected

$$|\partial_{v_i}\partial_{v_j}\mathcal{L}| = |\partial_{p_i}\partial_{p_j}\mathcal{H}|^{-1}. \tag{B.21}$$

That means that kinematic equivalency is exactly the invertibility condition of the Hessians [CA21]. Ergo only for systems with the kinematic equivalency condition, Lagrangian and Hamiltonian systems are equivalent. A connection between the formalisms is sketched in Fig. B.1.

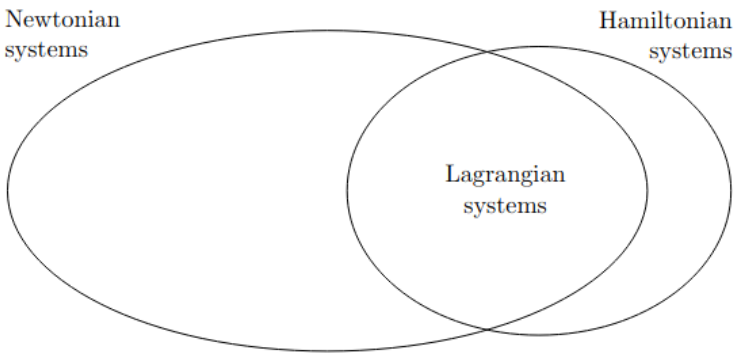


FIGURE B.1: Connection between the different formalisms of classical mechanics. They are not completely equivalent [CA21].

Appendix C

Elements of Stochastic Mechanics

C.1 Connection to other Variational Principles

The Hamilton principle

$$\delta S = \delta \int_{t_0}^{t_1} \mathcal{L}(t, q, \dot{q}) \mathrm{d}t = 0 \tag{C.1}$$

leads to the Euler-Lagrange equations, which gives a solution, if the matrix with elements

$$\frac{\partial^2 \mathcal{L}}{\partial \dot{x}_i \partial \dot{x}_j} \tag{C.2}$$

is non-negative definite. The introduction of a Hamilton function however only works, if the momentum and the time derivative of the coordinates are invertible,

$$p = p(x, \dot{x}, t), \quad \dot{x} = \dot{x}(x, p, t). \tag{C.3}$$

If this is not possible, one has to introduce a more general variational principle, the Pontryagin maximum principle. We may interpret this variational principle as an optimal control problem with the additional control

$$u(t) = \frac{\mathrm{d}x}{\mathrm{d}t}, \tag{C.4}$$

yielding a more general Hamiltonian

$$H(x,u,p,t) = pu - \mathcal{L}(x,u,t) \tag{C.5}$$

with a costate variable

$$p = \frac{\partial \mathcal{L}}{\partial u}. \tag{C.6}$$

The canonical equations

$$\frac{\partial H}{\partial p} = u = \frac{\mathrm{d}x}{\mathrm{d}t}, \tag{C.7}$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial x}, \tag{C.8}$$

are extended by

$$\frac{\partial H}{\partial u} = p - \frac{\partial \mathcal{L}}{\partial p} = 0 \tag{C.9}$$

with the matrix with elements

$$\frac{\partial^2 H}{\partial u_i \partial u_j} \tag{C.10}$$

being negative definite.
The formalism was used, adapted to stochastic processes, in deriving the quantum Hamilton equations.

C.2 Different Lagrangians

For the derivation of the quantum Hamilton equations in this thesis, the Lagrangian

$$\mathcal{L} = \frac{m}{2}(v - iu)^2 - V(x), \tag{C.11}$$

proposed by Michele Pavon [Pav95] in 1995 was used, combining the saddle point action principle and the saddle point entropy production principle.
Multiple Lagrangians have been proposed over the years. Kunio Yasue [Yas81a] proposed in 1981

$$\mathcal{L} = \frac{1}{2}\left(\frac{m}{2}(D_f x)^2 + \frac{m}{2}(D_b x)^2\right) - V(x) = \frac{m}{2}(v^2 + u^2) - V(x), \tag{C.12}$$

which he used to derive Nelson's stochastic, Newton's second law, and the gradient conditions of the velocity fields

$$u + v = \frac{\hbar}{m}\nabla S + \frac{\hbar}{2m}\nabla \ln \rho. \tag{C.13}$$

Two years later, Guerra and Morato [GM83] proposed

$$\mathcal{L} = \left(\frac{m}{2}(D_f x D_b x)\right) - V(x) = \frac{m}{2}(v^2 - u^2) - V(x), \tag{C.14}$$

which yielded the same results. However, due to the negative sign of the osmotic velocity, the connection to the quantum potential of Bohmian mechanics becomes evident. This Lagrangian is the real part, the saddle point action principle part, of Pavon's complex Lagrangian.

C.3 Stationary but Non-Ergodic Processes

Stationary processes are ergodic, if

$$\frac{1}{T}\int_0^T x_t \mathrm{d}t = \int x\rho(x)\mathrm{d}x. \tag{C.15}$$

The left side of this equation concerns the sample paths of a stochastic processes, but the right side concerns the probability density with

$$X_t \propto \rho(x). \tag{C.16}$$

This means that from a single sample path, if it is long enough, can deduce statistical properties of the process. From this follows that all ergodic processes are stationary. But not all stationary processes are ergodic.
Consider a one-dimensional quantum mechanical ground state of a symmetrical potential consisting of two wells that are separated by a wall of infinite height at position $x = 0$. This leads to a root in the probability distribution at position zero. From

$$u(x) \propto \partial_x \ln \rho(x) \tag{C.17}$$

one can see that the osmotic velocity diverges at $x = 0$. A stochastic trajectory that starts for example in the left well ($x < 0$) will never reach the other one at $x > 0$. Sampling that trajectory and calculating the mean will yield

$$\frac{1}{T}\int_0^T x_t \mathrm{d}t < 0 \qquad \forall T > 0 \tag{C.18}$$

Calculating the mean from the probability density will however yield

$$\int x\rho(x)\mathrm{d}x = 0 \tag{C.19}$$

due to the symmetry of the potential. $\rho(x)$ does not change over time and is thusly stationary in the physical sense. It is also stationary in the first kind due to $(X_1, \cdots, X_n) \overset{d}{=} (X_{1+h}, \ldots, X_{n+h})$. This is an example of an stationary process that is not ergodic.
All exited states are stationary but not ergodic, due to the nodes in the wave function. That is why the supersymmetric partner Hamiltonians were treated in sections 3.6 and 5.5.

Appendix D

The Riccati Equation

The Riccati equation is a non-linear differential equation of first order [Raj21; Hal97] with the general form

$$y' = a(x)y^2 + b(x)y + c(x). \tag{D.1}$$

The stationary Schrödinger equation

$$-\Psi''(x) = \frac{2m}{\hbar}(E - V(x))\Psi \tag{D.2}$$

can be transformed into a Riccati equation with the ansatz

$$W = -\frac{\Psi'}{\Psi}, \quad W' = -\frac{\Psi''}{\Psi} + \frac{\Psi'^2}{\Psi^2}. \tag{D.3}$$

With this

$$\Psi W' + \frac{\Psi'^2}{\Psi} = \frac{2m}{\hbar}(E - V(x))\frac{\Psi'}{W} \tag{D.4}$$

results in

$$W' = W^2 + \frac{2m}{\hbar}(E - V(x)). \tag{D.5}$$

From this equation, one can construct the ground states and the exited states [Raj21]. The stationary Schrödinger equation is a special case of a Sturm-Liouville equation

$$\alpha(x)u''(x) + \beta(x)u(x)' + \gamma(x)u(x) = 0. \tag{D.6}$$

A solution $u(x)$ of this ordinary differential equation also solves the Riccatti equation

$$y' + y^2 + \frac{\beta y + \gamma}{\alpha} = 0 \tag{D.7}$$

for $y = \frac{u'}{u}$ [Whe04].

Appendix E

Other Applications of the Algorithm

The algorithm discussed in chapter 5.7 was used, after its first publication in [HP23], for the solution of other kinds of quantum mechanical problems

E.1 Systems with Radial Symmetry

In Daniel Rodau. "Modellierung der Protonenemission von Thulium-145 mittels der Stochastischen Quantenmechanik." Bachelor's Thesis. Martin-Luther-University Halle-Wittenberg, 2024, the algorithm applied to systems with radial symmetry, first the hydrogen atom with potential

$$V(r) = -\frac{1}{4\pi\varepsilon_0}\frac{e^2}{r}. \tag{E.1}$$

and the Woods-Saxon potential

$$V(r) = -\frac{V_0}{1 + e^{\frac{r-R_0}{a}}}. \tag{E.2}$$

The osmotic velocity of the hydrogen atom is consistent with the previous attempts of

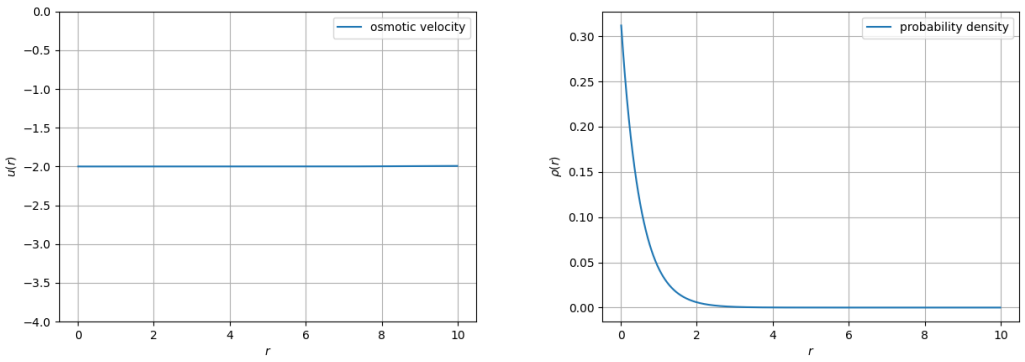


FIGURE E.1: Osmotic velocity and corresponding probability density of the ground state of the hydrogen atom [Rod24].

calculating the osmotic velocity with the Bender/Steiner algorithm [Kö+18].
The Woods-Saxon potential is an effective potential for a nucleon (a proton in this case) in a heavy nucleus with $V_0 = 50MeV$ being the depth of the potential, $R_0 = 1.25fm\sqrt[3]{A}$ being

the radius of the nucleus with the mass number A and $a \approx 0.5fm$ being the thickness of the nucleon wall, was calculated.

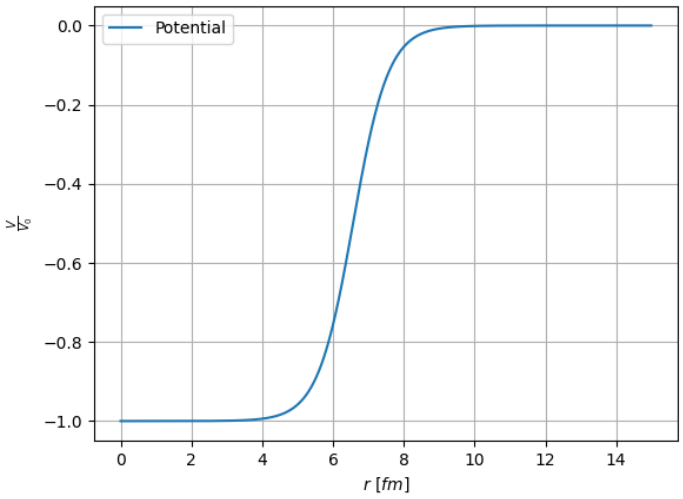


FIGURE E.2: Woods-Saxon potential [Rod24]

The resulting osmotic velocity was used to approximate the half-life for the proton emission of Thulium-145.

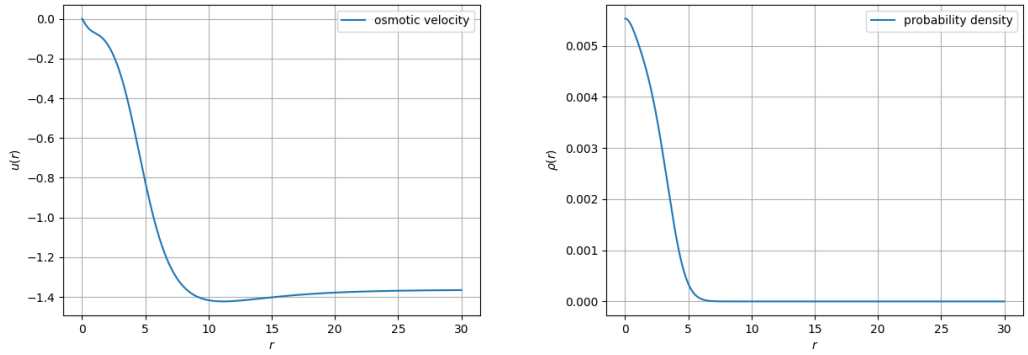


FIGURE E.3: Osmotic velocity and corresponding probability density of the ground state of the proton in a nucleus of Thulium-145 [Rod24].

E.2 Periodic Potentials - Rotators

In Paul Reißmann. "Eigenschaften der Methylrotation mittels der Stochastischen Mechanik." Bachelor's Thesis. Martin-Luther-University Halle-Wittenberg, 2025, the algorithm was used to describe a periodic rotator potential, introduced in [Kle24], to investigate the spectroscopic characteristics of molecules with a rotating methyl group (like acetaldehyde CH_3CHO molecule). The potential is given by

$$V(\varphi) = \frac{V_3}{2}(1 - \cos(3\varphi)) \tag{E.3}$$

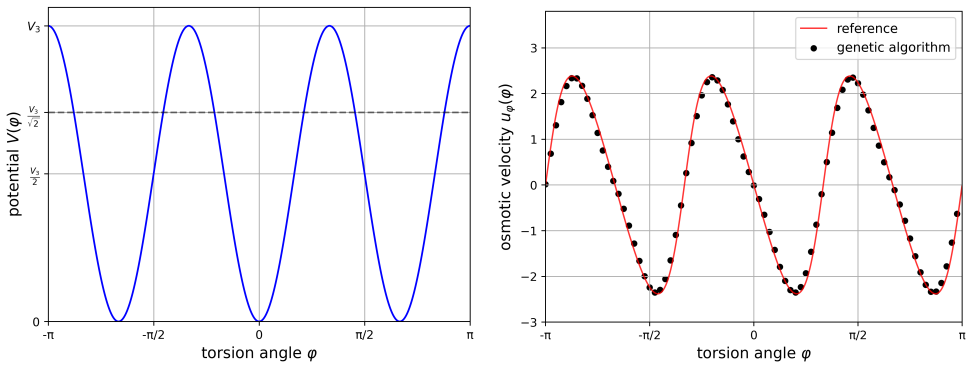


FIGURE E.4: Potential and corresponding osmotic velocity of the rotating methyl group of the molecule [Rei25].

with the index indicating the $\frac{2\pi}{3}$ periodicity of the potential. Higher orders like V_6 were not treated in [Rei25]. The potential and the osmotic velocity can be seen in Fig.(E.4). The osmotic velocity was used to calculate the time needed in the mean for the methyl group to rotate from one minimum to another in the same framework like one would calculate the tunneling time.

Bibliography

[AB40] Luis W. Alvarez and F. Bloch. "A Quantitative Determination of the Neutron Moment in Absolute Nuclear Magnetons." In: *Phys. Rev.* 57 (2 1940), pp. 111–122. DOI: 10.1103/PhysRev.57.111. URL: https://link.aps.org/doi/10.1103/PhysRev.57.111.

[AGY23] Naseem Alsadi, S. Andrew Gadsden, and John Yawney. "Intelligent estimation: A review of theory, applications, and recent advances." In: *Digital Signal Processing* 135 (2023), p. 103966. ISSN: 1051-2004. DOI: https://doi.org/10.1016/j.dsp.2023.103966. URL: https://www.sciencedirect.com/science/article/pii/S1051200423000611.

[Alo+22] Tahani A. Aloafi et al. "Predication and Photon Statistics of a Three-Level System in the Photon Added Negative Binomial Distribution." In: *Symmetry* 14.2 (2022). ISSN: 2073-8994. DOI: 10.3390/sym14020284. URL: https://www.mdpi.com/2073-8994/14/2/284.

[Arn73] L. Arnold. *Stochastische Differentialgleichungen: Theorie u. Anwendung*. Oldenbourg, 1973. ISBN: 9783486339413. URL: https://books.google.de/books?id=01jQAAAAMAAJ.

[Aug24] Midhun T Augustine. *A Survey on Universal Approximation Theorems*. 2024. arXiv: 2407.12895 [cs.LG]. URL: https://arxiv.org/abs/2407.12895.

[BAL70] L. E. BALLENTINE. "The Statistical Interpretation of Quantum Mechanics." In: *Rev. Mod. Phys.* 42 (4 1970), pp. 358–381. DOI: 10.1103/RevModPhys.42.358. URL: https://link.aps.org/doi/10.1103/RevModPhys.42.358.

[Bau23] Veronika Baumann. "Classical Information and Collapse in Wigner's Friend Setups." In: *Entropy* 25.10 (2023). ISSN: 1099-4300. DOI: 10.3390/e25101420. URL: https://www.mdpi.com/1099-4300/25/10/1420.

[BD16] Peter J. Brockwell and Richard A. Davis. *Introduction to Time Series and Forecasting*. Springer Texts in Statistics. Springer International Publishing, 2016. DOI: 10.1007/978-3-319-29854-2. URL: https://doi.org/10.1007/978-3-319-29854-2.

[BD86] Peter J Brockwell and Richard A Davis. *Time series: theory and methods*. Berlin, Heidelberg: Springer-Verlag, 1986. ISBN: 0387964061.

[Bee+19] Chris Beeler et al. "Optimizing thermodynamic trajectories using evolutionary reinforcement learning." In: *CoRR* abs/1903.08543 (2019). arXiv: 1903.08543. URL: http://arxiv.org/abs/1903.08543.

[Bel64] J. S. Bell. "On the Einstein Podolsky Rosen paradox." In: *Physics Physique Fizika* 1 (3 1964), pp. 195–200. DOI: 10.1103/PhysicsPhysiqueFizika.1.195. URL: https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195.

[Bey18] Michael Beyer. "DERIVATION OF QUANTUM HAMILTON EQUATIONS FOR MULTI-DIMENSIONAL SYSTEMS." MA thesis. Martin-Luther-University Halle-Wittenberg, 2018.

[Bey+19] M. Beyer et al. "Quantum Hamilton equations for multidimensional systems." In: *J. Phys. A* 52.16 (2019), p. 165301. DOI: 10.1088/1751-8121/ab0bcf.

[Bey23] Michael Beyer. "Particle spin described by quantum Hamilton equations." PhD thesis. 2023. URL: http://dx.doi.org/10.25673/112607.

[BG10] Seid Bahlali and Boulekhrass Gherbal. "Optimality conditions of controlled backward doubly stochastic differential equations." In: *Random Operators and Stochastic Equations* 18 (Sept. 2010). DOI: 10.1515/ROSE.2010.014.

[BJ25] Max Born and Pascual Jordan. "Zur Quantenmechanik." In: *Zeitschrift für Physik* 34 (1925), pp. 858–888. URL: https://api.semanticscholar.org/CorpusID:186114542.

[Bla67] John M Blatt. "Practical points concerning the solution of the Schrödinger equation." In: *Journal of Computational Physics* 1.3 (1967), pp. 382–396. ISSN: 0021-9991. DOI: https://doi.org/10.1016/0021-9991(67)90046-0. URL: https://www.sciencedirect.com/science/article/pii/0021999167900460.

[Bog+23] Alina Bogoi et al. "Assessment of Stochastic Numerical Schemes for Stochastic Differential Equations with "White Noise" Using Itô's Integral." In: *Symmetry* 15.11 (2023). ISSN: 2073-8994. DOI: 10.3390/sym15112038. URL: https://www.mdpi.com/2073-8994/15/11/2038.

[Boh05] N. Bohr. *On the Quantum Theory of Line-spectra*. Dover books on physics. Dover Publications, 2005. ISBN: 9780486442488. URL: https://books.google.de/books?id=MuydXvJ3uJMC.

[Boh13a] N. Bohr. "I. On the constitution of atoms and molecules." In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 26.151 (1913), pp. 1–25. DOI: 10.1080/14786441308634955. URL: https://doi.org/10.1080/14786441308634955.

[Boh13b] N. Bohr. "LXXIII. On the constitution of atoms and molecules." In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 26.155 (1913), pp. 857–875. DOI: 10.1080/14786441308635031. URL: https://doi.org/10.1080/14786441308635031.

[Boh13c] N. Bohr. "XXXVII. On the constitution of atoms and molecules." In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 26.153 (1913), pp. 476–502. DOI: 10.1080/14786441308634993. URL: https://doi.org/10.1080/14786441308634993.

[Bor26] Max Born. "Zur Quantenmechanik der Stosvorgänge." In: *Zeitschrift für Physik* 37.12 (1926), pp. 863–867.

[BP23] Michael Beyer and Wolfgang Paul. "Particle Spin Described by Quantum Hamilton Equations." In: *Annalen der Physik* 535.1 (2023), p. 2200433. DOI: https://doi.org/10.1002/andp.202200433. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.202200433. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.202200433.

[BPC16] A. Bacci, V. Petrillo, and M. Rossetti Conti. "GIOTTO: A Genetic Code for Demanding Beam-dynamics Optimizations." In: *Proc. of International Particle Accelerator Conference (IPAC'16), Busan, Korea, May 8-13, 2016* (Busan, Korea). International Particle Accelerator Conference 7. doi:10.18429/JACoW-IPAC2016-WEPOY039. Geneva, Switzerland: JACoW, 2016, pp. 3073–3076. ISBN: 978-3-95450-147-2. DOI: doi:10.18429/JACoW-IPAC2016-WEPOY039. URL: http://jacow.org/ipac2016/papers/wepoy039.pdf.

[Bro24] Louis de Broglie. "Recherches sur la théorie des Quanta." Theses. Migration - université en cours d'affectation, Nov. 1924. URL: https://theses.hal.science/tel-00006807.

[BS12] Christian Bender and Jessica Steiner. "Least-Squares Monte Carlo for Backward SDEs." In: *Numerical Methods in Finance*. Ed. by René A. Carmona et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 257–289. ISBN: 978-3-642-25746-9.

[Bur77] P. G. Burke. "Variational Methods and Bound Principles." In: *Potential Scattering in Atomic Physics*. Boston, MA: Springer US, 1977, pp. 75–96. ISBN: 978-1-4613-4112-3. DOI: 10.1007/978-1-4613-4112-3_8. URL: https://doi.org/10.1007/978-1-4613-4112-3_8.

[CA21] Gabriele Carcassi and Christine A. Aidala. *Assumptions of Physics*. Michigan Publishing, 2021. ISBN: 978-1-60785-706-7. DOI: 10.3998/mpub.12204707. URL: https://doi.org/10.3998/mpub.12204707.

[Car25] James Briggs Laura Carnevali. *AlexNet and ImageNet: The Birth of Deep Learning*. 2025. URL: https://www.pinecone.io/learn/series/image-search/imagenet/ (visited on 02/26/2025).

[CFZ16] Thomas Curtright, David Fairlie, and Cosmas Zachos. *A Concise Treatise on Quantum Mechanics in Phase Space*. Dec. 2016. ISBN: ISBN-13: 978-9814520430.

[CHH02] Murray Campbell, A.Joseph Hoane, and Feng hsiung Hsu. "Deep Blue." In: *Artificial Intelligence* 134.1 (2002), pp. 57–83. ISSN: 0004-3702. DOI: https://doi.org/10.1016/S0004-3702(01)00129-1. URL: https://www.sciencedirect.com/science/article/pii/S0004370201001291.

[Com20] Wikimedia Commons. *File:Gamma-ray-microscope.svg — Wikimedia Commons, the free media repository*. [Online; accessed 24-June-2025]. 2020. URL: https://commons.wikimedia.org/w/index.php?title=File:Gamma-ray-microscope.svg&oldid=451184083.

[Com24] Wikimedia Commons. *File:Optical trap principle formula edit.svg — Wikimedia Commons, the free media repository*. [Online; accessed 24-June-2025]. 2024. URL: https://commons.wikimedia.org/w/index.php?title=File:Optical_trap_principle_formula_edit.svg&oldid=858113384.

[CZ12] Thomas Curtright and Cosmas Zachos. "Quantum Mechanics in Phase Space." In: *Asia-Pacific Physics Newsletter* 1 (Jan. 2012), pp. 37–46. DOI: 10.1142/S2251158X12000069.

[DEK60] A. Dvoretzky, P. Erdős, and S. Kakutani. "Nonincrease Everywhere of the Brownian Motion Process." In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability (June 20-July 30 1960)* (1960). Referenced by: MathSciNet [MR0132608], pages 103–116. URL: http://digicoll.lib.berkeley.edu/record/112926.

[DF26] Paul Adrien Maurice Dirac and Ralph Howard Fowler. "On the theory of quantum mechanics." In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 112.762 (1926), pp. 661–677. DOI: 10.1098/rspa.1926.0133. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1926.0133. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1926.0133.

[DG27] C. Davisson and L. H. Germer. "Diffraction of Electrons by a Crystal of Nickel." In: *Phys. Rev.* 30 (6 1927), pp. 705–740. DOI: 10.1103/PhysRev.30.705. URL: https://link.aps.org/doi/10.1103/PhysRev.30.705.

[Dio87] L. Diosi. "A Universal Master Equation for the Gravitational Violation of Quantum Mechanics." In: *Phys. Lett. A* 120 (1987), p. 377. DOI: 10.1016/0375-9601(87)90681-5.

[Dir81] P.A.M. Dirac. *The Principles of Quantum Mechanics*. Comparative Pathobiology - Studies in the Postmodern Theory of Education. Clarendon Press, 1981. ISBN: 9780198520115. URL: https://books.google.de/books?id=XehUpGiM6FIC.

[DM89] N. David Mermin. "What's Wrong with this Pillow?" In: *Physics Today* 42.4 (Apr. 1989), pp. 9–11. ISSN: 0031-9228. DOI: 10.1063/1.2810963. eprint: https://pubs.aip.org/physicstoday/article-pdf/42/4/9/8301150/9_1_online.pdf. URL: https://doi.org/10.1063/1.2810963.

[EH27] Kennard E. H. "Zur Quantenmechanik einfacher Bewegungstypen." In: *Zeitschrift für Physik* 44 (1927), pp. 326–352. DOI: https://doi.org/10.1007/BF01391200.

[Ein05] A. Einstein. "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt." In: *Annalen der Physik* 322.6 (1905), pp. 132–148. DOI: https://doi.org/10.1002/andp.19053220607. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19053220607. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19053220607.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" In: *Phys. Rev.* 47 (10 1935), pp. 777–780. DOI: 10.1103/PhysRev.47.777. URL: https://link.aps.org/doi/10.1103/PhysRev.47.777.

[Eve57] Hugh Everett. ""Relative State" Formulation of Quantum Mechanics." In: *Rev. Mod. Phys.* 29 (3 1957), pp. 454–462. DOI: 10.1103/RevModPhys.29.454. URL: https://link.aps.org/doi/10.1103/RevModPhys.29.454.

[FD22] Ulysse Faure and Florent Draye. "A Review on the Universal Approximation Theorems of Artificial Neural Networks." PhD thesis. July 2022. DOI: 10.13140/RG.2.2.14106.25287.

[FHS10] R.P. Feynman, A.R. Hibbs, and D.F. Styer. *Quantum Mechanics and Path Integrals*. Dover Books on Physics. Dover Publications, 2010. ISBN: 9780486477220. URL: https://books.google.de/books?id=JkMuDAAAQBAJ.

[Flü12] S. Flügge. *Practical Quantum Mechanics*. Classics in Mathematics. Springer Berlin Heidelberg, 2012. ISBN: 9783642619953. URL: https://books.google.de/books?id=Md5sCQAAQBAJ.

[Fra19] Giuseppe Franco. *Handbuch Karl Popper*. Jan. 2019. ISBN: 978-3-658-16238-2. DOI: 10.1007/978-3-658-16239-9.

[Gar09] C. Gardiner. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer Series in Synergetics. Springer Berlin Heidelberg, 2009. ISBN: 9783540707127. URL: https://books.google.de/books?id=otg3PQAACAAJ.

[Gar85] C.W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Proceedings in Life Sciences. Springer-Verlag, 1985. ISBN: 9783540113577. URL: https://books.google.de/books?id=cRfvAAAAMAAJ.

[GI18] Namig J. Guliyev and Vugar E. Ismailov. "On the approximation by single hidden layer feedforward neural networks with fixed weights." In: *Neural Networks* 98 (2018). ISSN: 0893-6080. DOI: 10.1016/j.neunet.2017.12.007. URL: http://dx.doi.org/10.1016/j.neunet.2017.12.007.

[Gla63] Roy J. Glauber. "Coherent and Incoherent States of the Radiation Field." In: *Phys. Rev.* 131 (6 1963), pp. 2766–2788. DOI: 10.1103/PhysRev.131.2766. URL: https://link.aps.org/doi/10.1103/PhysRev.131.2766.

[Gle18] Sergei Gleyzer. *The rise of deep learning*. 2018. URL: https://cerncourier.com/a/the-rise-of-deep-learning/ (visited on 02/26/2025).

[GLW05] Emmanuel Gobet, Jean-Philippe Lemor, and Xavier Warin. "A regression-based Monte Carlo method to solve backward stochastic differential equations." In: *The Annals of Applied Probability* 15.3 (2005), pp. 2172 –2202. DOI: 10.1214/105051605000000412. URL: https://doi.org/10.1214/105051605000000412.

[GM83] Francesco Guerra and Laura M. Morato. "Quantization of dynamical systems and stochastic control theory." In: *Phys. Rev. D* 27 (8 1983), pp. 1774–1786. DOI: 10.1103/PhysRevD.27.1774. URL: https://link.aps.org/doi/10.1103/PhysRevD.27.1774.

[Gol89] David E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. New York: Addison-Wesley, 1989.

[GS22] Wālther Gerlāch and Otto Stern. "Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld." In: *Zeitschrift für Physik* 9 (1922), pp. 349–352. URL: https://api.semanticscholar.org/CorpusID:121543218.

[GW01] A. Gallant and Halbert White. "There Exists A Neural Network That Does Not Make Avoidable Mistakes." In: *[No source information available]* (Apr. 2001).

[Haa21] D. de Haan. *Numerical time integration of stochastic differential equations*. 2021. URL: http://essay.utwente.nl/88493/.

[Hab07] Fathi Habashi. "Alvarez, Luis Walter." In: *The Biographical Encyclopedia of Astronomers*. Ed. by Thomas Hockey et al. New York, NY: Springer New York, 2007, pp. 38–39. ISBN: 978-0-387-30400-7. DOI: 10.1007/978-0-387-30400-7_41. URL: https://doi.org/10.1007/978-0-387-30400-7_41.

[Hal97] Stephen B. Haley. "An underrated entanglement: Riccati and Schrödinger equations." In: *American Journal of Physics* 65.3 (Mar. 1997), pp. 237–243. DOI: 10.1119/1.18535.

[Hea23] Richard Healey. "Quantum-Bayesian and Pragmatist Views of Quantum Theory." In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Winter 2023. Metaphysics Research Lab, Stanford University, 2023.

[Hei27] W. Heisenberg. "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik." In: *Zeitschrift für Physik* 43 (1927), pp. 172–198. DOI: https://doi.org/10.1007/BF01397280.

[Her35] Grete Hermann. "Die naturphilosophischen Grundlagen der Quantenmechanik." In: *Naturwissenschaften* 23.42 (1935), pp. 718–721.

[HH10] Ting Hu and Simon Harding. "Variable population size and evolution acceleration: A case study with a parallel evolutionary algorithm." In: *Genetic Programming and Evolvable Machines* 11 (June 2010), pp. 205–225. DOI: 10.1007/s10710-010-9105-2.

[Hil02] David Hilbert. "Mathematical problems." In: *Bulletin of the American Mathematical Society* 8.10 (1902), pp. 437 –479.

[HL99] Georges Harik and Fernando Lobo. "A parameter-less genetic algorithm." In: Jan. 1999, pp. 258–265.

[Hol62] John H. Holland. "Outline for a Logical Theory of Adaptive Systems." In: *J. ACM* 9 (1962), pp. 297–314. URL: https://api.semanticscholar.org/CorpusID:955154.

[Hol75] John H. Holland. *Adaptation in Natural and Artificial Systems*. second edition, 1992. Ann Arbor, MI: University of Michigan Press, 1975.

[Hop82] J J Hopfield. "Neural networks and physical systems with emergent collective computational abilities." In: *Proceedings of the National Academy of Sciences* 79.8 (1982), pp. 2554–2558. DOI: 10.1073/pnas.79.8.2554. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.79.8.2554. URL: https://www.pnas.org/doi/abs/10.1073/pnas.79.8.2554.

[Hop84] J J Hopfield. "Neurons with graded response have collective computational properties like those of two-state neurons." In: *Proceedings of the National Academy of Sciences* 81.10 (1984), pp. 3088–3092. DOI: 10.1073/pnas.81.10.3088. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.81.10.3088. URL: https://www.pnas.org/doi/abs/10.1073/pnas.81.10.3088.

[HP23] Kai-Hendrik Henk and Wolfgang Paul. "Machine learning quantum mechanical ground states based on stochastic mechanics." In: *Phys. Rev. A* 108 (6 2023), p. 062412. DOI: 10.1103/PhysRevA.108.062412. URL: https://link.aps.org/doi/10.1103/PhysRevA.108.062412.

[HP25] Kai-Hendrik Henk and Wolfgang Paul. *Levitodynamics: An analysis of quantum fluctuations based on stochastic mechanics*. 2025. arXiv: 2506.16459 [quant-ph]. URL: https://arxiv.org/abs/2506.16459.

[HT85] John Hopfield and D Tank. "Neural Computation of Decisions in Optimization Problems." In: *Biological cybernetics* 52 (Feb. 1985), pp. 141–52. DOI: 10.1007/BF00339943.

[IM88] Irie and Miyake. "Capabilities of three-layered perceptrons." In: *IEEE 1988 International Conference on Neural Networks*. 1988, 641–648 vol.1. DOI: 10.1109/ICNN.1988.23901.

[Itô44] Kiyosi Itô. "Stochastic integral." In: *Proceedings of the Imperial Academy* 20.8 (1944), pp. 519 –524. DOI: 10.3792/pia/1195572786. URL: https://doi.org/10.3792/pia/1195572786.

[JMJ21] Alexander Pritzel Tim Green Michael Figurnov Olaf Ronneberger Kathryn Tunyasuvunakool Russ Bates Augustin Žídek Anna Potapenko Alex Bridgland Clemens Meyer Simon A A Kohl Andy Ballard Andrew Cowie Bernardino Romera-Paredes Stanislav Nikolov Rishub Jain Jonas Adler Trevor Back Stig Petersen David Reiman Ellen Clancy Michal Zielinski Martin Steinegger Michalina Pacholska Tamas Berghammer Sebastian Bodenstein David Silver Oriol Vinyals Andrew W. Senior Koray Kavukcuoglu Pushmeet Kohli Demis Hassabis John M. Jumper Richard Evans. "Highly accurate protein structure prediction with AlphaFold." In: *Nature* 596 (2021), pp. 583 –589. URL: https://api.semanticscholar.org/CorpusID:235959867.

[Joh77] B. R. Johnson. "New numerical methods applied to solving the one-dimensional eigenvalue problem." In: *The Journal of Chemical Physics* 67.9 (Nov. 1977), pp. 4086–4093. ISSN: 0021-9606. DOI: 10.1063/1.435384. eprint: https://pubs.aip.org/aip/jcp/article-pdf/67/9/4086/18908788/4086_1_online.pdf. URL: https://doi.org/10.1063/1.435384.

[KCK20] Sourabh Katoch, Sumit Singh Chauhan, and Vijay Kumar. "A review on genetic algorithm: past, present, and future." In: *Multimedia Tools and Applications* 80 (2020), pp. 8091–8126. DOI: 10.1007/s11042-020-10139-6. URL: https://doi.org/10.1007/s11042-020-10139-6.

[Keu07] Herbert Keuth, ed. *Karl Popper: Logik der Forschung*. Berlin: Akademie Verlag, 2007. ISBN: 9783050050188. DOI: doi:10.1524/9783050050188. URL: https://doi.org/10.1524/9783050050188.

[KGP16] Jeanette Köppe, Wilfried Grecksch, and Wolfgang Paul. "Derivation and application of quantum Hamilton equations of motion." In: *Annalen der Physik* 529 (Dec. 2016), p. 1600251. DOI: 10.1002/andp.201600251.

[Khi34] A. Khintchine. "Korrelationstheorie der stationären stochastischen Prozesse." In: *Mathematische Annalen* 109 (1934), pp. 604–615. URL: https://api.semanticscholar.org/CorpusID:122842868.

[Kle24] Isabell Kleiner. "Spectroscopy of Molecules." In: *Physik Journal* 23 (2024), p. 63. eprint: https://pro-physik.de/zeitschriften/physik-journal/2024-9/. URL: https://pro-physik.de/zeitschriften/physik-journal/2024-9/.

[Klo02] Peter Kloeden. "The Systematic Derivation of Higher Order Numerical Schemes for Stochastic Differential Equations." In: *Milan Journal of Mathematics* 70 (Jan. 2002), pp. 187–207. DOI: 10.1007/s00032-002-0006-6.

[Kol57] A. N. Kolmogorov. "On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition." In: *Dokl. Akad. Nauk SSSR* 114 (1957), pp. 953–956. ISSN: 0002-3264.

[KP11] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2011. ISBN: 9783540540625. URL: https://books.google.de/books?id=BCvtssom1CMC.

[Kra40] H.A. Kramers. "Brownian motion in a field of force and the diffusion model of chemical reactions." In: *Physica* 7.4 (1940), pp. 284–304. ISSN: 0031-8914. DOI: https://doi.org/10.1016/S0031-8914(40)90098-2. URL: https://www.sciencedirect.com/science/article/pii/S0031891440900982.

[KS85] John Klaudert and Bo-Sture Skagerstam. *Coherent States: Applications in Physics and Mathematical Physics*. Jan. 1985. ISBN: 9971-966-52-2. DOI: 10.1142/0096.

[KSH12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "ImageNet Classification with Deep Convolutional Neural Networks." In: *Advances in Neural Information Processing Systems*. Ed. by F. Pereira et al. Vol. 25. Curran Associates, Inc., 2012. URL: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

[Kö18] Jeanette Köppe. "Derivation and Application of Quantum Hamilton Equations of Motion." PhD thesis. Feb. 2018.

[Kö+18] Jeanette Köppe et al. "Quantum Hamilton equations of motion for bound states of one-dimensional quantum systems." In: *Journal of Mathematical Physics* 59 (June 2018), p. 062102. DOI: 10.1063/1.5026377.

[Kö+20] Jeanette Köppe et al. "Quantum Hamilton Equations from Stochastic Optimal Control Theory." In: *Infinite Dimensional and Finite Dimensional Stochastic Equations and Applications in Physics*. May 2020, pp. 213–250. ISBN: 978-981-12-0978-9. DOI: 10.1142/9789811209796_0005.

[Kůr92] Věra Kůrková. "Kolmogorov's theorem and multilayer neural networks." In: *Neural Networks* 5 (1992), pp. 501–506. URL: https://api.semanticscholar.org/CorpusID:5748809.

[LB21] Matteo Lostaglio and Joseph Bowles. "The original Wigner's friend paradox within a realist toy model." In: *Proceedings of the Royal Society A* 477 (2021). URL: https://api.semanticscholar.org/CorpusID:239028682.

[LBL39] Frédéric London, Edmond Bauer, and Paul Langevin. "La théorie de l'observation en mécanique quantique." In: 1939. URL: https://api.semanticscholar.org/CorpusID:56536199.

[LD24] Olimpia Lombardi and Dennis Dieks. "Modal Interpretations of Quantum Mechanics." In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Summer 2024. Metaphysics Research Lab, Stanford University, 2024.

[LF87] Alan Lapedes and Robert Farber. "How Neural Nets Work." In: Jan. 1987, pp. 442–456. ISBN: 978-9971-5-0529-5. DOI: 10.1142/9789814434102_0012.

[LTE02] P.S. de Laplace, F.W. Truscott, and F.L. Emory. *A Philosophical Essay on Probabilities*. A Philosophical Essay on Probabilities. Wiley, 1902. URL: https://books.google.de/books?id=WxoPAAAAIAAJ.

[Lu+17] Zhou Lu et al. "The Expressive Power of Neural Networks: A View from the Width." In: (Sept. 2017). DOI: 10.48550/arXiv.1709.02540.

[Mad26] Erwin Madelung. "Eine anschauliche Deutung der Gleichung von Schrödinger." In: *Naturwissenschaften* 14 (1926), p. 1004. DOI: https://doi.org/10.1007/BF01504657.

[Mag+21] Lorenzo Magrini et al. "Real-time optimal quantum control of mechanical motion at room temperature." In: *Nature* 595.7867 (July 2021), pp. 373–377. ISSN: 1476-4687. DOI: 10.1038/s41586-021-03602-3. URL: http://dx.doi.org/10.1038/s41586-021-03602-3.

[Max65] James Clerk Maxwell. "VIII. A dynamical theory of the electromagnetic field." In: *Philosophical Transactions of the Royal Society of London* 155 (1865), pp. 459–512. DOI: 10.1098/rstl.1865.0008. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rstl.1865.0008. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rstl.1865.0008.

[MD89] David J. Montana and Lawrence Davis. "Training Feedforward Neural Networks Using Genetic Algorithms." In: *International Joint Conference on Artificial Intelligence*. 1989. URL: https://api.semanticscholar.org/CorpusID:6336712.

[MGS24] Wayne Myrvold, Marco Genovese, and Abner Shimony. "Bell's Theorem." In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Spring 2024. Metaphysics Research Lab, Stanford University, 2024.

[Mil10] R. A. Millikan. "The Isolation of an Ion, a Precision Measurement of Its Charge, and the Correction of Stokes's Law." In: *Science* 32.822 (1910), pp. 436–448. DOI: 10.1126/science.32.822.436. eprint: https://www.science.org/doi/pdf/10.1126/science.32.822.436. URL: https://www.science.org/doi/abs/10.1126/science.32.822.436.

[Mil75] G. N. Mil'shtejn. "Approximate Integration of Stochastic Differential Equations." In: *Theory of Probability & Its Applications* 19.3 (1975), pp. 557–562. DOI: 10.1137/1119062. eprint: https://doi.org/10.1137/1119062. URL: https://doi.org/10.1137/1119062.

[Mit98] Melanie Mitchell. *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press, 1998. ISBN: 0262631857.

[Nel66] Edward Nelson. "Derivation of the Schrödinger Equation from Newtonian Mechanics." In: *Phys. Rev.* 150 (4 1966), pp. 1079–1085. DOI: 10.1103/PhysRev.150.1079. URL: https://link.aps.org/doi/10.1103/PhysRev.150.1079.

[Nel85] Edward Nelson. *Quantum Fluctuations*. Princeton University Press, 1985. ISBN: 9780691083797. URL: http://www.jstor.org/stable/j.ctv14163xm (visited on 12/10/2024).

[Nol13] Wolfgang Nolting. *Grundkurs: Theoretische Physik 5/1 Quantenmechanik-Grundlagen*. ger. 8. Aufl. Berlin: Springer, 2013. ISBN: 3540421122.

[Nol14] Wolfgang Nolting. *Grundkurs: Theoretische Physik 2 Analytische Mechanik*. ger. 9. Aufl. Berlin: Springer, 2014. ISBN: 3540421122.

[Nor17] T. Norsen. *Foundations of Quantum Mechanics: An Exploration of the Physical Meaning of Quantum Theory*. Undergraduate Lecture Notes in Physics. Springer International Publishing, 2017. ISBN: 9783319658674. URL: https://books.google.de/books?id=BaQxDwAAQBAJ.

[Øks10] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Universitext. Springer Berlin Heidelberg, 2010. ISBN: 9783642143946. URL: https://books.google.de/books?id=EQZEAAAAQBAJ.

[ØS14] Bernt Øksendal and Agnès Sulem. "Forward–Backward Stochastic Differential Games and Stochastic Control under Model Uncertainty." In: *Journal of Optimization Theory and Applications* 161 (Apr. 2014). DOI: 10.1007/s10957-012-0166-7.

[Par+20] Sejun Park et al. *Minimum Width for Universal Approximation*. 2020. arXiv: 2006.08859 [cs.LG]. URL: https://arxiv.org/abs/2006.08859.

[Pat18] Markus Patzold. "Excited States in Nelson's Stochastic Mechanics." MA thesis. Martin-Luther-University Halle-Wittenberg, 2018.

[Pav00] Michele Pavon. "Stochastic mechanics and the Feynman integral." In: *Journal of Mathematical Physics* 41.9 (Sept. 2000), pp. 6060–6078. ISSN: 0022-2488. DOI: 10.1063/1.1286880. eprint: https://pubs.aip.org/aip/jmp/article-pdf/41/9/6060/19275376/6060_1_online.pdf. URL: https://doi.org/10.1063/1.1286880.

[Pav95] Michele Pavon. "Hamilton's principle in stochastic mechanics." In: *Journal of Mathematical Physics* 36 (Dec. 1995), pp. 6774–6800. DOI: 10.1063/1.531187.

[Pav99] Michele Pavon. "Derivation of the wave function collapse in the context of Nelson's stochastic mechanics." In: *Journal of Mathematical Physics* 40.11 (Nov. 1999), pp. 5565–5577. ISSN: 0022-2488. DOI: 10.1063/1.533046. eprint: https://pubs.aip.org/aip/jmp/article-pdf/40/11/5565/19014191/5565_1_online.pdf. URL: https://doi.org/10.1063/1.533046.

[PB13] W. Paul and J. Baschnagel. *Stochastic Processes: From Physics to Finance*. Second Edition. Springer International Publishing, 2013. ISBN: 9783319003276. URL: https://books.google.de/books?id=OWANAAAAQBAJ.

[PE97] Hans Primas and Michael Esfeld. "A Critical Review of Wigner's Work on the Conceptual Foundations of Quantum Theory." In: (Sept. 1997).

[Pen96] Roger Penrose. "On gravity's role in quantum state reduction." In: *Gen. Rel. Grav.* 28 (1996), pp. 581–600. DOI: 10.1007/BF02105068.

[Per77] A M Perelomov. "Generalized coherent states and some of their applications." In: *Soviet Physics Uspekhi* 20.9 (1977), p. 703. DOI: 10.1070/PU1977v020n09ABEH005459. URL: https://dx.doi.org/10.1070/PU1977v020n09ABEH005459.

[Pla01] Max Planck. "Ueber das Gesetz der Energieverteilung im Normalspectrum." In: *Annalen der Physik* 309.3 (1901), pp. 553–563. DOI: https://doi.org/10.1002/andp.19013090310. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19013090310. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19013090310.

[PT33] G. Pöschl and Edward Teller. "Bemerkungen zur Quantenmechanik des anharmonischen Oszillators." In: *Zeitschrift für Physik* 83 (1933), pp. 143–151. URL: https://api.semanticscholar.org/CorpusID:124830271.

[Qui06] R.S. Quimby. "Appendix B: Fourier Synthesis and the Uncertainty Principle." In: *Photonics and Lasers*. John Wiley Sons, Ltd, 2006, pp. 499–504. ISBN: 9780471791591. DOI: https://doi.org/10.1002/0471791598.app2.

[R09] Andreas Rößler. "Second Order Runge–Kutta Methods for Itô Stochastic Differential Equations." In: *SIAM Journal on Numerical Analysis* 47.3 (2009), pp. 1713–1738. DOI: 10.1137/060673308. eprint: https://doi.org/10.1137/060673308. URL: https://doi.org/10.1137/060673308.

[R10] Andreas Rößler. "Runge–Kutta Methods for the Strong Approximation of Solutions of Stochastic Differential Equations." In: *SIAM Journal on Numerical Analysis* 48.3 (2010), pp. 922–952. DOI: 10.1137/09076636X. eprint: https://doi.org/10.1137/09076636X. URL: https://doi.org/10.1137/09076636X.

[Raj21] Kazimierz Rajchel. "A New Constructive Method for Solving the Schrödinger Equation." In: *Symmetry* 13.10 (2021). ISSN: 2073-8994. DOI: 10.3390/sym13101879. URL: https://www.mdpi.com/2073-8994/13/10/1879.

[Red17] Maik Reddiger. "The Madelung Picture as a Foundation of Geometric Quantum Theory." In: *Foundations of Physics* 47.10 (Sept. 2017), pp. 1317–1367. ISSN: 1572-9516. DOI: 10.1007/s10701-017-0112-5. URL: http://dx.doi.org/10.1007/s10701-017-0112-5.

[Rei25] Paul Reißmann. "Eigenschaften der Methylrotation mittels der Stochastischen Mechanik." Bachelor's Thesis. Martin-Luther-University Halle-Wittenberg, 2025.

[Rel22] Press Release. *2022 Nobel Prize Laureates in Physics: Alain Aspect, John F. Clauser, and Anton Zeilinger*. 2022. URL: https://www.nobelprize.org/prizes/physics/2022/summary/ (visited on 02/27/2025).

[Rel24a] Press Release. *2024 Nobel Prize Laureates in Chemistry: David Baker, Demis Hassabis and John Jumper*. 2024. URL: https://www.nobelprize.org/prizes/physics/2024/press-release/ (visited on 02/27/2025).

[Rel24b] Press Release. *2024 Nobel Prize Laureates in Physics: John J. Hopfield and Geoffrey Hinton*. 2024. URL: https://www.nobelprize.org/prizes/physics/2024/press-release/ (visited on 02/27/2025).

[RF12] H. Risken and T. Frank. *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer Series in Synergetics. Springer Berlin Heidelberg, 2012. ISBN: 9783642615443. URL: https://books.google.de/books?id=6B5uCQAAQBAJ.

[RHW86] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors." In: *Nature* 323 (1986), pp. 533–536. URL: https://api.semanticscholar.org/CorpusID:205001834.

[Rit09] Walter Ritz. "Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik." ger. In: *Journal für die reine und angewandte Mathematik* 135 (1909), pp. 1–61. URL: http://eudml.org/doc/149295.

[Rod24] Daniel Rodau. "Modellierung der Protonenemission von Thulium-145 mittels der Stochastischen Quantenmechanik." Bachelor's Thesis. Martin-Luther-University Halle-Wittenberg, 2024.

[Rov25] Carlo Rovelli. "Relational Quantum Mechanics." In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Spring 2025. Metaphysics Research Lab, Stanford University, 2025.

[Rut12] E. Rutherford. "The scattering of and particles by matter and the structure of the atom." In: *Philosophical Magazine* 92.4 (2012), pp. 379–398. DOI: 10.1080/14786435.2011.617037. eprint: https://doi.org/10.1080/14786435.2011.617037. URL: https://doi.org/10.1080/14786435.2011.617037.

[San18] Shabnam Sangwan. "Literature Review on Genetic Algorithm." In: *International Journal of Research* 5 (June 2018), p. 1142.

[SB20] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. 2nd ed. Cambridge, Massachusetts: The MIT Press, 2020.

[Sch10] Thomas Schaffter. *Numerical Integration of SDEs: A Short Tutorial*. École Polytechnique Fédérale de Lausanne (EPFL), 2010. URL: https://infoscience.epfl.ch/handle/20.500.14299/45604.

[Sch26a] Erwin Schrödinger. "Der stetige Übergang von der Mikro- zur Makromechanik." In: *Naturwissenschaften* 14 (1926), pp. 664–666. URL: https://api.semanticscholar.org/CorpusID:34680073.

[Sch26b] E. Schrödinger. "Quantisierung als Eigenwertproblem." In: *Annalen der Physik* 79.18 (1926), pp. 109–139. DOI: https://doi.org/10.1002/andp.19263861802.

[Sch26c] E. Schrödinger. "Quantisierung als Eigenwertproblem." In: *Annalen der Physik* 79.6 (1926), pp. 489–527. DOI: https://doi.org/10.1002/andp.19263840602.

[Sch26d] E. Schrödinger. "Quantisierung als Eigenwertproblem." In: *Annalen der Physik* 80.13 (1926), pp. 437–490. DOI: https://doi.org/10.1002/andp.19263851302.

[Sch26e] E. Schrödinger. "Quantisierung als Eigenwertproblem." In: *Annalen der Physik* 81.18 (1926), pp. 109–139. DOI: https://doi.org/10.1002/andp.19263861802.

[Sch26f] Erwin Schrödinger. "Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinem." In: *Annalen der Physik* 384.8 (1926), pp. 734–756. DOI: https://doi.org/10.1002/andp.19263840804.

[Sch99] Werner Scholz. *Micromagnetic Simulation of Thermally Activated Switching in Fine Particles*. Diploma thesis. 1999. URL: http://www.cwscholz.net/projects/da/da.html.

[Sci22] The Royal Swedisch Academy of Science. *Entangled states – from theory to technology*. 2022. URL: https://www.nobelprize.org/prizes/physics/2022/press-release/.

[See16] Michiel Seevinck. "Challenging the Gospel: Grete Hermann on von Neumann's No-Hidden-Variables Proof." In: *Grete Hermann - Between Physics and Philosophy*. Ed. by Elise Crull and Guido Bacciagaluppi. Dordrecht: Springer Netherlands, 2016, pp. 107–117. ISBN: 978-94-024-0970-3. DOI: 10.1007/978-94-024-0970-3_7. URL: https://doi.org/10.1007/978-94-024-0970-3_7.

[SP10] Shambhu Sharma and Hiren Patel. "The Fokker-Planck Equation." In: Aug. 2010. ISBN: 978-953-307-121-3. DOI: 10.5772/9730.

[Spe00] William Spears. *Evolutionary Algorithms*. Jan. 2000. ISBN: 978-3-642-08624-3. DOI: 10.1007/978-3-662-04199-4.

[Teg07] Max Tegmark. "Many lives in many worlds." In: *Nature* 448.7149 (July 2007), pp. 23–24. ISSN: 1476-4687. DOI: 10.1038/448023a. URL: http://dx.doi.org/10.1038/448023a.

[Tho95] J. J. Thomson. "XLVIII. The relation between the atom and the charge of electricity carried by it." In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 40.247 (1895), pp. 511–544. DOI: 10.1080/14786449508620801. URL: https://doi.org/10.1080/14786449508620801.

[Tho98] J. J. Thomson. "LVII. On the charge of electricity carried by the ions produced by Röntgen rays." In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 46.283 (1898), pp. 528–545. DOI: 10.1080/14786449808621229. URL: https://doi.org/10.1080/14786449808621229.

[TUR50] A. M. TURING. "I.—COMPUTING MACHINERY AND INTELLIGENCE." In: *Mind* LIX.236 (Oct. 1950), pp. 433–460. ISSN: 0026-4423. DOI: 10.1093/mind/LIX.236.433. eprint: https://academic.oup.com/mind/article-pdf/LIX/236/433/61209000/mind_lix_236_433.pdf. URL: https://doi.org/10.1093/mind/LIX.236.433.

[VJ23] Amala Vincent and P. Jidesh. "An improved hyperparameter optimization framework for AutoML systems using evolutionary algorithms." In: *Scientific Reports* 13 (Mar. 2023). DOI: 10.1038/s41598-023-32027-3.

[Von32] John VonNeumann. *Mathematische Grundlagen der Quantenmechanik*. ger. Berlin [u.a.]: Springer, 1932. URL: http://eudml.org/doc/203794.

[Wal94] Timothy C. Wallstrom. "Inequivalence between the Schrödinger equation and the Madelung hydrodynamic equations." In: *Phys. Rev. A* 49 (3 1994), pp. 1613–1617. DOI: 10.1103/PhysRevA.49.1613. URL: https://link.aps.org/doi/10.1103/PhysRevA.49.1613.

[Wel67] P. Welch. "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms." In: *IEEE Transactions on Audio and Electroacoustics* 15.2 (1967), pp. 70–73. DOI: 10.1109/TAU.1967.1161901.

[Whe04] Nicholas Wheeler. "Quantum Applications of the Riccati Equation." In: *Reed College Physics Department* (June 2004). URL: https://www.reed.edu/physics/faculty/wheeler/documents/Quantum%20Mechanics/Miscellaneous%20Essays/Riccati%20Method%20in%20QM.pdf.

[Wig95] E. P. Wigner. "Remarks on the Mind-Body Question." In: *Philosophical Reflections and Syntheses*. Ed. by Jagdish Mehra. Berlin, Heidelberg: Springer Berlin Heidelberg, 1995, pp. 247–260. ISBN: 978-3-642-78374-6. DOI: 10.1007/978-3-642-78374-6_20. URL: https://doi.org/10.1007/978-3-642-78374-6_20.

[Wil03] F. Williams. *Topics in Quantum Mechanics*. Progress in Mathematical Physics. Birkhäuser Boston, 2003. ISBN: 9780817643119. URL: https://books.google.de/books?id=lcc1BNA7NroC.

[WM67] Eugene Wigner and Henry Margenau. "Symmetries and Reflections, Scientific Essays." In: *American Journal of Physics* 35.12 (Dec. 1967), pp. 1169–1170. ISSN: 0002-9505. DOI: 10.1119/1.1973829. eprint: https://pubs.aip.org/aapt/ajp/article-pdf/35/12/1169/11891421/1169_1_online.pdf. URL: https://doi.org/10.1119/1.1973829.

[Yas80] Kunio Yasue. "Schrödinger's variational method of quantization revisited." In: *Letters in Mathematical Physics* 4 (1980), pp. 143–146. DOI: https://doi.org/10.1007/BF00417507.

[Yas81a] Kunio Yasue. "Quantum mechanics and stochastic control theory." In: *Journal of Mathematical Physics* 22.5 (May 1981), pp. 1010–1020. ISSN: 0022-2488. DOI: 10.1063/1.525006. eprint: https://pubs.aip.org/aip/jmp/article-pdf/22/5/1010/19178211/1010_1_online.pdf. URL: https://doi.org/10.1063/1.525006.

[Yas81b] Kunio Yasue. "Stochastic calculus of variations." In: *Journal of Functional Analysis* 41.3 (1981), pp. 327–340. ISSN: 0022-1236. DOI: https://doi.org/10.1016/0022-1236(81)90079-3. URL: https://www.sciencedirect.com/science/article/pii/0022123681900793.

[Yas83] Kunio Yasue. "A variational principle for the Navier-Stokes equation." In: *Journal of Functional Analysis* 51.2 (1983), pp. 133–141. ISSN: 0022-1236. DOI: https://doi.org/10.1016/0022-1236(83)90021-6. URL: https://www.sciencedirect.com/science/article/pii/0022123683900216.

Curriculum Vitae

Personal Data

Name: Henk, Kai-Hendrik

Education

since May 2021 **PhD student**, *Martin-Luther-Universität*, Halle Wittenberg, *Theoretical Polymer Physics*
Supervisor: Prof. Wolfgang Paul

Sep. 2020 **Master of Science, Physics**, *Martin-Luther-Universität*, Halle Wittenberg, *Theoretical Polymer Physics*
Supervisor: PD. Jan Kantelhardt

Sep. 2018 **Bachelor of Science, Physics**, *Martin-Luther-Universität*, Halle Wittenberg, *Mesoscopic Quantum Dynamics*
Supervisor: Prof. Miguel Marques

Jul. 2014 **Abitur**, *Elisabeth Gymnasium*, Halle(Saale)

Signed: _____

Date: _____

List of publications

Kai-Hendrik Henk and Wolfgang Paul. “Machine learning quantum mechanical ground states based on stochastic mechanics.” In: *Phys. Rev. A* 108 (6 2023), p. 062412. DOI: 10.1103/PhysRevA.108.062412. URL: https://link.aps.org/doi/10.1103/PhysRevA.108.062412

Declaration of Authorship

I, Kai-Hendrik Henk, hereby declare that the work presented in this thesis is my own and has been written under the supervision of, Prof. Dr. Wolfgang Paul. I did not use any outside support except for the quoted literature and other sources mentioned in the manuscript.

Signed:

Date:

Acknowlegment

I would like to thank all the people that made the completion of this thesis possible.

First of all, I would like to thank and express my deep gratitude to Prof. Dr. Wolfgang Paul for giving me the opportunity to work on this project. During my time in his research group, he was always open for discussions and his insights, not only related to physics, taught me a lot. His guidance and patience helped me tremendously. My time in his research group was a joy and I am grateful for his enduring support.
I would also like to thank Dr. Viktor Ivanov for helping me ease into my role as a teacher and for being always a friendly and helpful presence in the group.
I would like to thank my colleagues Lama Tannoury, Evgeniia Filimonova, Michael Beyer, Christian Sommerfeld, Timur Shakirov and Prof. Steffen Trimper for making the time truly memorable. Thank you especially for the many discussions, not always scientific, during our lunch and coffee breaks.
I am privileged to have been able to spend this time with you.

I would like to thank Theresa Berthold, Daniel Rodau, Paul Reißmann and Christian Dancker for their work during their bachelor's theses and Philipp Tesch for his work on his bachelor's and master's theses. A supervisor could not ask for better students.

Special thanks goes out to Philine, Antonia, Friederike, Esther, Jan and Juli. Our weekly lunch breaks were always a joy and your continuing presence since our time together during our bachelor studies helped me a lot to grow as a researcher and as a person.

I'd like to extend my gratitude to my friends and family. Without your continuing love and support, all of this would not have been possible. I would like to especially thank those who had to endure me the most during the stressful part of my work. Words cannot describe how much gratitude I owe each and everyone of you.

thank you all for everything!