Different Approaches to Stochastic Quantisation in One-Dimensional Processes

bachelor's thesis

Submitted to the Faculty of Mathematics, Computer Science and Natural Sciences at RWTH Aachen University

presented by

Leona Rodenkirchen

under the supervision of

Prof. Dr. Fabian Hassler

JARA Institute for Quantum Information

08/2020

Abstract

This thesis examines three different approaches to stochastic quantisation in one dimension: stochastic differential equations, path integrals and Fokker-Planck equations. The issue of non-differentiability of a stochastic process is resolved by discretising time in increments $\varepsilon > 0$. *We find that all three approaches are equivalent up to arbitrary accuracy in that they produce the same results in the continuum limit* $\varepsilon \to 0$. *This is non-trivial as they not only make different assumptions on the physical model but also show very different levels of abstraction; from an intuitive dynamic model given by stochastic differential equations over a generalisation of the action principle in path integrals to a Schrödinger-like Fokker-Planck equation that completely leaves the notion of trajectories behind and can be solved analytically. We thus conclude that classical motivations can be the basis for semiclassical models and the impairment of the accuracy of their predictions is negligible.*

Eidesstattliche Versicherung

Contents

Chapter 1

Introduction

1.1 Introduction and chapter overview

Quantum mechanics is a probabilistic theory and thereby closely intertwined with stochastics by its very nature. In physics, stochastic quantisation typically refers to a specific method developed by the physicists Giorgio Parisi and Yong Shi Wu in 1981 as a means to constructing quantum mechanics and quantum field theory as "the thermal equilibrium limit of a hypothetical stochastic process" [1]. For many problems in quantum field theory, the Parisi-Wu method has proven advantageous over the conventional theories [2]. However, this thesis does not aim at investigating the Parisi-Wu stochastic quantisation method but rather explores the merits and challenges that arise from stochastic processes and constructing a theory based on stochastic quantisation in general. Quantum theory is only one example discipline that can be modelled with stochastic processes. The Black-Scholes model for the financial-market dynamic is a renowned mathematical model in economics based on stochastic processes. Another important application is modelling a diffusive system in thermodynamics.

In the course of this thesis, we take a close look at three different methods to formulate stochastic processes in one dimension, that is stochastic differential equations, stochastic path integrals, and Fokker-Planck equations. There is a chapter dedicated to each of these three approaches. Chapter 2 focuses on stochastic differential equations (SDE) and introduces the procedure of discretising time. We discuss the requirements for a dynamic model of a system influenced by random forces, particularly the requirement of convergence in the continuum limit, that is the limit of infinitesimal discrete-time increments. In Sec. 2.2 and Sec. 2.3, we examine the SDEs of two important example pro-

cesses, Brownian motion and geometric Brownian motion. At the end of Ch.2, we analyse a general SDE and the process resulting from an arbitrary variable transformation performed on a general SDE. This analysis results in fundamental differences between SDEs and normal non-stochastic differential equations, including a non-unique discrete representation and the violation of the chain rule of differential calculus.

In Ch.3, we construct stochastic path integrals from the probability density functions of random variable sequences defined by a discrete stochastic process. While the SDE approach allows an understanding of the microscopic dynamic of a process, the path-integral approach considers entire paths as single events and delivers the probability of possible trajectories of a stochastic process. In Sec.3.2 of this chapter, we introduce the Hubbard-Stratonovich transformation, which is then used in Sec.3.3 for performing a variable transformation in the path integral.

The third method, namely the Fokker-Planck equation, is discussed in Ch.4. In Sec.4.1, we derive the Fokker-Planck equation of a stochastic process based on the path integral results from Ch.3. However, the resulting formalism works independent from path integrals. The Fokker-Planck equation is a non-stochastic second order partial differential equation that reduces a stochastic process to only its starting point and the probability to arrive in a specific end point. We find that it shows close resemblance to the Schrödinger equation. Other than for the other two approaches, the calculations in Fokker-Planck equations are entirely analytic. We see in Sec.4.2 that this makes variable transformations a lot easier for Fokker-Planck equations.

In this thesis, we find many similarities between stochastic processes and the concepts of quantum mechanics. This can be seen as a motivation to approach quantum mechanical problems as questions of stochastic processes. Section 5.2 of the conclusion offers a closer look into the possibilities arising from such an approach to quantum mechanics.

1.2 Random walk

The fundamental idea of a stochastic process is well captured by the concept of a random walk. It describes an object ‘walking’ randomly driven by a force that is stochastically distributed; thus, neither direction nor strength of the

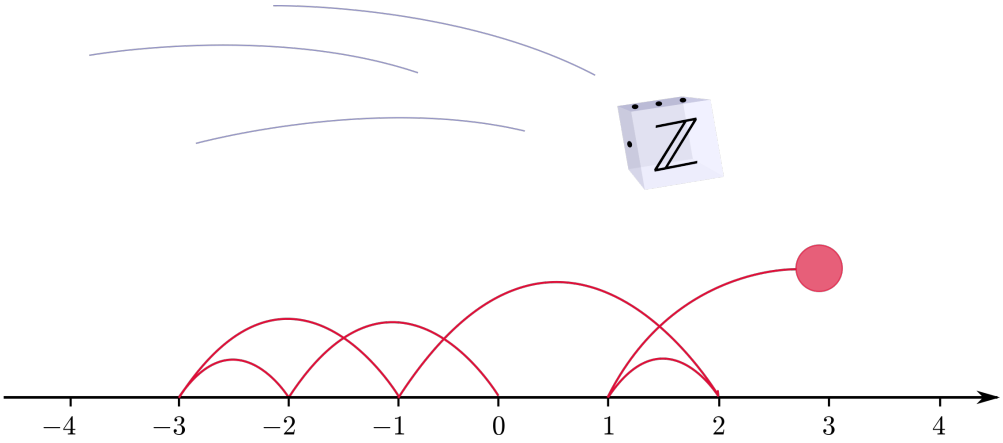


Figure 1.1: Illustration of an object undertaking a random walk in the set of all integers starting at zero. The object changes positions each time a specific amount of time passes. The distance travelled with each step is determined by a so-called random force, which gives out a random integer number, much like throwing dice. The outcome of a random walk after a certain amount of steps cannot be reproduced due to the random character of the motion.

impact are determined. The outcome of such a walk cannot be reproduced from the same starting conditions. It is stochastically distributed. We can only attempt at making predictions based on probabilities. Figure 1.1 illustrates the behaviour of an object undertaking a random walk in the set of integers. We figuratively 'throw the dice' on every step.

The mathematical model of a random walk is fairly simple, especially in one dimension. An object, for example a particle, moves back and forth randomly along a straight line. At a time t_0 it starts at some point x_0. After each time interval $\varepsilon > 0$ it jumps back or forth by $\Delta x(t) := x(t+\varepsilon) - x(t)$. The distance travelled with each step is determined by a function $F[x(t), \xi(t)]$. This function is called *random force*. It depends linearly on $\{\xi(t),\ t \in \mathbb{R}\}$, a set of independent and identically distributed stochastic variables called *white noise*. The set $\{x(t),\ t \in \mathcal{R}\}$ is called a *stochastic process*. For further elaboration on the mathematical background in probability theory see App. A. The behaviour of a random walk can be modelled by

$$\frac{\Delta x(t)}{\varepsilon} = F[x(t), \xi(t)]. \tag{1.1}$$

In the limit $\varepsilon \to 0$, Eq. (1.1) is called a *stochastic differential equation* (SDE), due to the stochastic character of the random force.

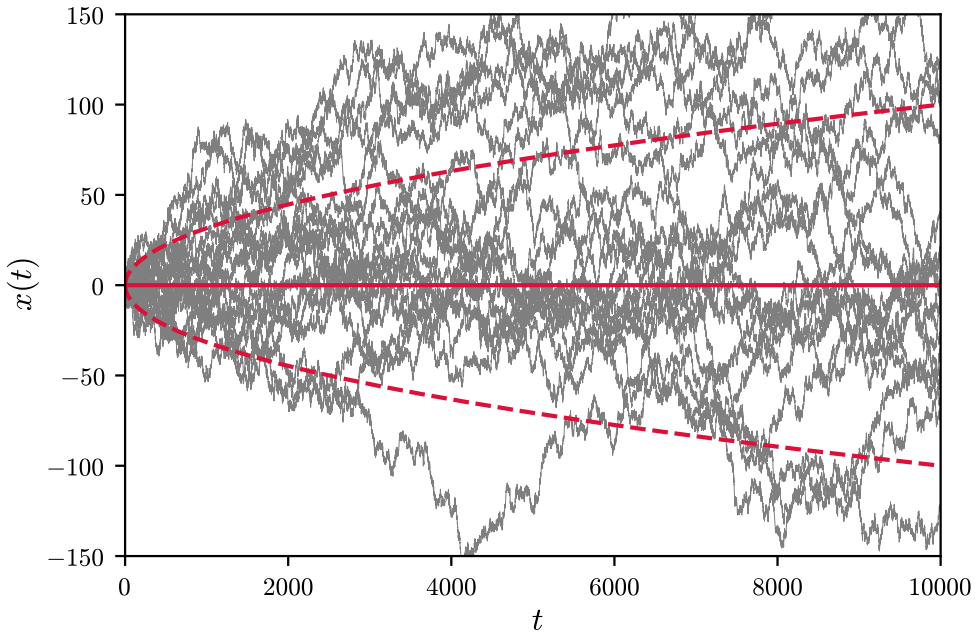


Figure 1.2: Displacement $x(t)$ of twenty example paths for discrete one-dimensional Brownian motion over time t with time increments of $\varepsilon = 0.1$, starting position of $x(0) = 0$ within a total time of $T = 10000$. The dashed red line displays the standard deviation of the path distribution growing with the square-root of the time. The continuous red line marks the mean position which is constant at $x(0) = 0$.

Even if we know exactly how the white noise $\xi(t)$ is distributed for all times, it is impossible to predict exactly the position $x(t)$ of one specific particle at $t = j\varepsilon$, $j \in \mathbb{N}$. However, we are capable of predicting with great accuracy the average behaviour of a large number of particles or of many trial runs with one particle, particularly their mean position $\langle x(t) \rangle$ and their variance $\langle\langle x(t)^2 \rangle\rangle := \langle x(t)^2 \rangle - \langle x(t) \rangle^2$. Dependent on the form of the SDE, it is even possible to determine the specific distribution of $x(t)$ in some cases.

An example of a process for which the determination of the specific path distribution from the SDE is feasible is the Wiener process named after the American mathematician Norbert Wiener. This is the stochastic process that models one-dimensional Brownian motion, which is thoroughly discussed in Sec. 2.2. In a Wiener process, the random force is given by $F[\xi(t)] = \xi(t)$, for all $t \in \mathbb{R}$. This corresponds to the simplest form of a SDE, that is $\Delta x(t) = \varepsilon \xi(t)$. For all its simplicity, it has a broad range of applications and the analysis of this model offers valuable insight into the general requirements for modelling a stochastic process. We learn in Sec. 2.2 that the position $x(t)$ of a Brownian motion path

is normally distributed for all $t \in \mathbb{R}$ with a constant mean value and a variance growing linear in t. Figure 1.2 shows a few example paths of Brownian motion as well as its characteristic constant mean and square-root standard deviation.

Chapter 2

Stochastic differential equations

2.1 Discrete time and continuum limit

In this thesis, we examine the behaviour of functions that are not differentiable in the classical analytical sense. Nevertheless, we can assign them derivative equivalents. In stochastic calculus, the issue of the non-differentiability of stochastic processes is overcome by working with integral equations in the place of differential equations. In physics, we make use of quantisation instead. We discretise time. In discrete time, a differential equation becomes a difference equation so that the problem of non-differentiability does not arise explicitly. We perform all the calculations in discrete time and project the results onto continuous reality by conducting the continuous-time limit called *continuum limit*. This procedure is what we call *stochastic quantisation*. In order to understand how this is realised and where things deviate from the usual, we briefly revisit some basics of analysis in this section.

The derivative of a function is a measure of its instantaneous rate of change in a specific point. As such, it can be defined as the approximative limit of the function's inclination within the immediate environment surrounding the point of interest. Mathematically, this is expressed in the *differential quotient* as the limit of the *difference quotient*. A function $f : V \subset \mathbb{R} \to \mathbb{R}$ is *differentiable*, if the limit

$$\left.\frac{\partial f(x)}{\partial x}\right|_{x=x_0} := \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \tag{2.1}$$

exists for all $x_0 \in V$ [3]. We are interested in the behaviour of paths $x : [0, \infty] \to \mathbb{R}$, $t \mapsto x(t)$ that identify each point in time with a position in one-dimensional real space. The difference quotients of such paths as given by Eq. (1.1) are the main interest of this chapter. They represent the change of the position of $x(t)$ in discrete steps $\varepsilon > 0$ of time. We examine discrete steps $\varepsilon = t/j$, $j \in \mathbb{N}$, and the stochastic distribution of the transition to the next position. Although, the model represents a continuous-time reality and thus only the results for $\varepsilon \to 0$ ($t = \text{const.}$) are relevant. This limit is known as the *continuum limit*. In this thesis, only terms that do not vanish in the continuum limit are of interest.

Another reason to evaluate stochastic differential equations in discrete time is that they are not unique in the continuum limit. This ambiguity stems from the random force defining a stochastic process. Given a first order differential equation for a path such as

$$\frac{\partial x(t)}{\partial t} = F[x(t)], \tag{2.2}$$

the discrete-time interpretation of the left side is delivered by Eq. (2.1) since the derivative is the difference quotient's limit for infinitesimal steps in time. However, it is unclear, how the right side of Eq. (2.2) needs to be interpreted for a finite step $\varepsilon > 0$. It can be $F[x(t)]$ as much as $F[x(t+\varepsilon)]$ or an arbitrary superposition of the possible interpretations as in

$$\frac{x(t+\varepsilon) - x(t)}{\varepsilon} = \alpha F[x(t+\varepsilon)] + (1-\alpha)F[x(t)], \tag{2.3}$$

with $\alpha \in [0, 1]$. The choice does not matter as long as it delivers the same result for $\varepsilon \to 0$. For a differentiable path, the choice of α does not make a difference in the continuum limit. The mean value theorem delivers the exact point $s \in [x(t), x(t+\varepsilon)]$ at which Eq. (2.3) holds true [3]. In this case, the difference between the value of F at that point and the value at any other point in the interval $[x(t), x(t+\varepsilon)]$ is of the order of ε and hence vanishes in the continuum limit, that is $F[\alpha x(t+\varepsilon) + (1-\alpha)x(t)] - F(s) = \mathcal{O}(\varepsilon)$ for all $\alpha \in [0, 1]$. Stochastic processes do not fulfil the requirement of differentiability for this argument. For SDEs, the right-side function F depends on a random variable $\xi(t)$, the white noise, in a linear manner. We see in Sec. 2.2 that the magnitude of $\xi(t)$ is dependent on the increment ε. Therefore, the choice of α that is made in discrete time generally makes a difference for stochastic processes in the continuum limit. This leads to an ambiguity of the differential equations particular to stochastic processes if no additional information on the interpretation of α

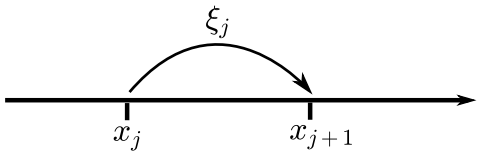
is provided. An in-depth discussion of this phenomenon follows in Sec. 2.4. In stochastic calculus, the debate over the proper interpretation of the right side of an SDE kick-started two competing schools of thought. An *Itō SDE* implies $\alpha = 0$. It is named after the Japanese mathematician Itō Kiyoshi. The other popular choice is $\alpha = \frac{1}{2}$. Such an SDE is called *Stratonovich SDE* after the Russian physicist Ruslan Stratonovich [4]. Section 2.4 and Sec. 2.5 further analyse the Itō vs. Stratonovich debate, particularly the benefits of both interpretations. For now, the point is to become familiar with the notion that a continuous differential equation is not unique for a stochastic process and performing calculations in discrete time is a way to avoid this obstacle.

Before diving into the general case, there are two concrete examples worth considering first. These examples are meant to illustrate the characteristics of the white noise $\xi(t)$ and point out some peculiar effects that come with random forces.

2.2 Brownian motion

The first example is the Brownian motion already introduced briefly in Sec. 1.2. Brownian motion describes the random movement of particles in a fluid caused by collisions with fluid molecules. In 1905, Albert Einstein famously developed a model for Brownian motion based on identifying effectively stochastic forces causing the diffusive motion and additional drift forces [5]. Shortly after this breakthrough, the French physicist Paul Langevin streamlined the mathematical framework for Brownian motion introducing the Langevin equation, an equation of motion for a Brownian motion particle. The Langevin equation also includes deterministic forces which cause a drift and stochastic forces which model the diffusive behaviour of the system. Einstein's and Langevin's approaches are both rooted in the concept of stochastic quantisation which indicates the discontinuous structure of matter causing Brownian motion in the first place [1].

We limit ourselves to stochastic forces only in this discussion of the Brownian motion in order to examine the characteristics of the white noise. The analysis of this Wiener-process model for Brownian motion is going to reveal that the standard deviation of $\xi(t)$, $t \in \mathbb{R}$, needs to scale with ε in order to guarantee the convergence of the model in the continuum limit. That is a principle which lies at the basis of modelling stochastic processes and the main focus of this section.



$$\langle x_j \xi_j \rangle = 0$$
$$\langle x_{j+1} \xi_j \rangle \neq 0$$

Figure 2.1: Chronology and correlation in discrete Brownian motion. In Brownian motion, the displacement of a path $(x_j)_j \subset \mathbb{R}$ at each step $x_j \to x_{j+1}$ is directly given by the white noise value ξ_j. The variables x_j and ξ_j are not correlated. Only the position following thereafter is in direct correlation with ξ_j indicating the displayed causality.

Furthermore, we are going to see that we can assume a normal distribution of $\xi(t)$ for all $t \in \mathbb{T}$ without loss of generality. All other processes discussed in this thesis are based on the results of this model of Brownian motion. The continuous SDE of the one-dimensional Brownian motion can be modelled by

$$\frac{\partial x(t)}{\partial t} = \xi(t),\ t \in \mathbb{R}, \tag{2.4}$$

with the following specifications for the white noise variable. We assume there is no deterministic drift inherent in the white noise and therefore choose $\langle \xi(t) \rangle = 0$ for all $t \in \mathbb{R}$. Furthermore, we demand a constant-in-time standard deviation $\sigma(t) \equiv \sigma$ and no correlation due to chronology, such that $\langle \xi(t)\xi(t+\varepsilon) \rangle = 0$ for all $t \in \mathbb{R}$. The latter ensures that $\{\xi(t),\ t \in \mathbb{R}\}$ are independent and identically distributed variables. In the case of Eq. (2.4), the random force given by the right side of the equation does not depend on x. Therefore, the question of α, as in how to interpret the x-value on the right side, is redundant for Brownian motion. Without further ado, we can write the SDE in discrete time as

$$\boxed{\frac{x_{j+1} - x_j}{\varepsilon} = \xi_j,} \tag{2.5}$$

where $j \in \mathbb{N}$ is the counting integer of time, such that $t = j\varepsilon$. Figure 1.2 displays the path behaviour for an example of Brownian motion.

We do not vary the right side of Eq. (2.5) explicitly in t, for example by using $\beta\xi_{j+1} + (1-\beta)\xi_j$, for $\beta \in \mathbb{R}$. The reason for leaving this possibility unaddressed is the inner logic of the model. We demand that we 'roll the dice' *after* the last step and *before* the next. This means that ξ_j and x_{j+1} are correlated, while x_j and ξ_j are not. Figure 2.1 offers a visualisation of the chronology. This subtlety is the reason that the derivative $\partial x/\partial t$ and the position x do not commute in

the continuum limit. We have $\langle \xi_j \rangle = 0$ and $\langle \xi_j^2 \rangle = \langle\langle \xi_j^2 \rangle\rangle = \sigma^2$ for all $j \in \mathbb{N}$. That yields

$$\begin{aligned} \lim_{\varepsilon \to 0} \langle \dot{x}(t+\varepsilon) x(t) \rangle &= \lim_{\varepsilon \to 0} \langle \xi_j \rangle \langle x_j \rangle = 0, \\ \lim_{\varepsilon \to 0} \langle x(t) \dot{x}(t-\varepsilon) &= \lim_{\varepsilon \to 0} \langle x_j \xi_{j-1} \rangle \neq 0. \end{aligned} \tag{2.6}$$

The exact value of the latter is calculated later in this section. Either way, this reminds of quantum mechanics, where we famously have $\langle px \rangle \neq \langle xp \rangle$. The order of a product indicates the chronology in which the values are measured. In physical reality, the two quantities cannot be evaluated at the exact same time; one must come first, the other second. Equation (2.6) proves that the order in which this is done makes a difference.

The mean position $\langle x(t) \rangle$ and the standard deviation $\langle\langle x(t)^2 \rangle\rangle$ of a path $x(t)$ for $t \in \mathbb{R}$ can be directly derived from the microscopic dynamic of the process defined by the discrete SDE. Equation (2.5) produces an expression for x_{j+1} in terms of x_j from which we can calculate the mean

$$\langle x_{j+1} \rangle = \langle x_j \rangle + \varepsilon \underbrace{\langle \xi_j \rangle}_{=0}. \tag{2.7}$$

This leads to a continuous mean that is constant over time, $\langle x(t) \rangle = x_0$. The constant x_0, which is the starting point, can be chosen to be zero without loss of generality, since it only represents a global translation. The standard deviation, which describes the average distance travelled up to t, can be addressed likewise. A recursive calculation yields

$$x_{j+1} = \underbrace{x_0}_{\equiv 0} + \varepsilon \sum_{i=0}^{j} \xi_i, \tag{2.8}$$

an expression of x_{j+1} as a function of $(\xi_j)_j$ which results in $\langle x_j^2 \rangle = \langle\langle x_j^2 \rangle\rangle = j\varepsilon^2\sigma^2$. If the standard deviation σ were constant in ε, that would mean $\langle\langle x_j^2 \rangle\rangle$ vanishes in the continuum limit and our model would thereby fail to reproduce the behaviour of a diffusive system as is the Brownian motion. This impediment can be resolved by scaling σ with ε. With $t = j\varepsilon$, we obtain $\langle\langle x(t)^2 \rangle\rangle = t\varepsilon\sigma^2$ which converges to a finite value in the continuum limit if and only if $\sigma^2 \propto 1/\varepsilon$. We set $D \in \mathbb{R}\backslash\{0\}$ as the proportionality constant alluding to the *diffusion coefficient* or *diffusivity* of thermodynamics. This yields an average distance

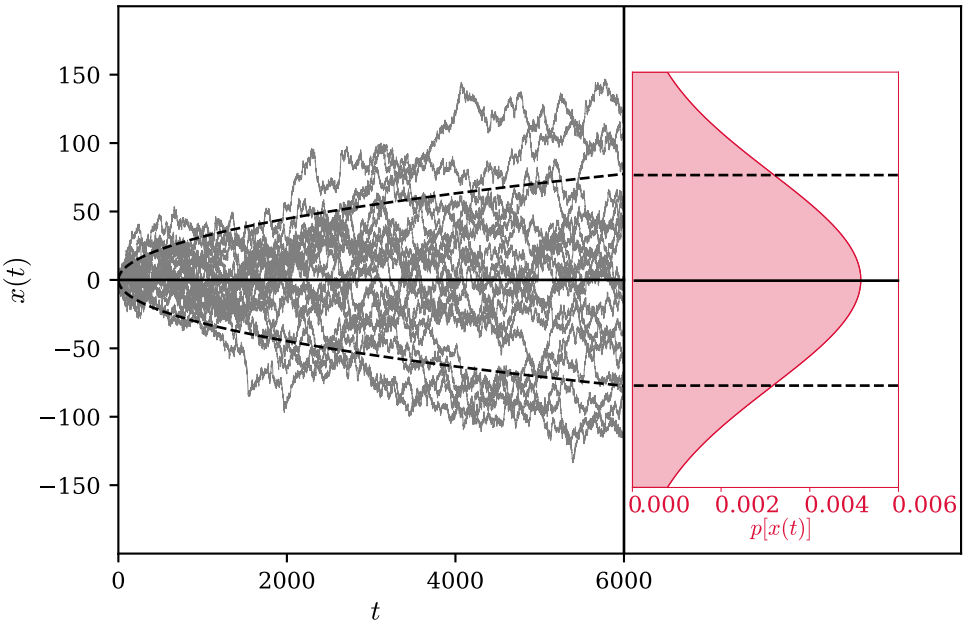


Figure 2.2: Gaussian distribution of the displacement of Brownian motion paths after a total time of $T = 6000$ in increments of $\varepsilon = 0.1$ starting at $x(0) = 0$ with a proportionality constant of $D = 1$. The dashed black line shows the standard deviation of the path distribution increasing with the square-root of time, while the continuous black line displays the constant mean value at $x(0) = 0$. The red plot on the right shows the ideal distribution of the paths at T given by a Gaussian probability distribution with a mean value of $\mu = 0$ and a standard deviation of $\sigma = \sqrt{6000}$.

travelled proportional to the square root of the time passed given by

$$\boxed{\sqrt{\langle\langle x(t)^2\rangle\rangle} = \sqrt{Dt}.} \tag{2.9}$$

These results prove that the inequality statement from Eq. (2.6) holds true. If Eq. (2.8) is inserted into $\langle x_j \xi_{j-1}\rangle$, we obtain

$$\langle x(t)\dot{x}(t)\rangle = \lim_{\varepsilon\to 0} \underbrace{\langle x_0 \xi_{j-1}\rangle}_{=0} + \varepsilon \underbrace{\langle \xi_{j-1}^2\rangle}_{=D/\varepsilon} = D, \tag{2.10}$$

making use of the fact that $(\xi_j)_j \subset \mathbb{R}$ are independent and identically distributed variables with a standard deviation antiproportional to ε.

Brownian motion paths follow a Gaussian probability distribution for all t. Therefore, the mean value and the standard deviation contain all information on the path distribution. This becomes apparent in Eq. (2.8). The central limit

theorem states that $\sum_{i=0}^{N} \xi_i$ is normally distributed for large N no matter the specific distribution of the random variables $\xi(t)$, $t \in \mathbb{R}$. For a set time $T = N\varepsilon$, the continuum limit is equivalent to the limit $N \to \infty$. Figure 2.2 shows a few exemplary Brownian motion paths whose collective behaviour shows indications of the process' Gaussian distribution which becomes apparent for a large number of paths.

The Gaussian character of Brownian motion is independent of the specific distribution of the white noise, not only due to the central limit theorem. For all processes in this thesis, the distribution of $\xi(t)$ can be assumed to be Gaussian for all $t \in \mathbb{R}$. All moments of the white noise of higher order than $\langle \xi(t)^2 \rangle$ are suppressed in the calculations for moments of the stochastic process $x(t)$. Since the random is always linear in $\xi(t)$ and the SDE yields $\Delta x = \varepsilon F[x(t), \xi(t)]$, the n-th white noise moment is multiplied and thus suppressed with ε^n. The moments uniquely define a distribution. In the continuum limit, the moments of $x(t)$ only depend on mean and variance of $\xi(t)$ for all $t \in \mathbb{R}$. We conclude that there are no constraints on the choice of the ξ-distribution for the model aside from mean value and standard deviation. Therefore, we can assume without loss of generality that $\xi(t)$ follows a Gaussian distribution for all t.

The process-specific proportionality constant D is often explicitly included in the SDE as in

$$\frac{\partial x(t)}{\partial t} = \sqrt{D}\xi(t). \tag{2.11}$$

Then, the variance of the white noise is $\sigma^2 = 1/\varepsilon$. All other results remain the same. For the sake of consistency, the random variable $\xi(t)$ in this thesis always has the same characteristics. Therefore, all proportionality constants like D are explicitly included in the SDE from here on.

It has been mentioned before that stochastic processes mostly define non-differentiable functions. In fact, it can be shown that the paths of a Wiener process are almost surely not differentiable in any point [6]. This result is called the Paley-Wiener-Zygmund theorem. The mathematical details behind this are not the matter of this thesis, but it shall not go unmentioned since we have already seen effects of the non-differentiability of the Wiener process in the few observations we have made so far.

2.3 Geometric Brownian motion

The second important example process is the *geometric Brownian motion* (GBM). This is the process a stock market follows in the Black-Scholes model [7]. Its defining characteristic is the percental, non-constant growth of the random force. Geometric Brownian motion is a good example to show the consequences of choosing an Itō SDE in a variable transformation. We are going to see that GBM follows a *log-normal* distribution. That means that the logarithm of its paths is normally distributed. Upon examining its logarithmic transform, we are going to see that the chain rule of differential calculus is violated.

The continuous SDE for GBM can be modelled by

$$\frac{\partial x(t)}{\partial t} = \sqrt{D} x(t) \xi(t),\ t \in \mathbb{R}, \tag{2.12}$$

with the same normally distributed white noise $\xi(t)$ as introduced in the previous section. For simplicity, additional deterministic forces are not considered here. The consideration of deterministic effects is part of the broad analysis in the subsequent section on general stochastic processes. The same goes for a general discussion of α. In this section, we assume $\alpha = 0$ in order to examine Itō-specific effects as a consequence of the variable transformation later in this section. The Itō SDE in discrete time is given by

$$\boxed{\frac{x_{j+1} - x_j}{\varepsilon} = \sqrt{D} x_j \xi_j.} \tag{2.13}$$

Same as in the previous section on Brownian motion, x_{j+1} can be expressed in terms of x_j through Eq. (2.13). Since x_j and ξ_j are uncorrelated, the mean value

$$\langle x_{j+1} \rangle = \langle x_j \rangle + \varepsilon \sqrt{D} \underbrace{\langle \xi_j x_j \rangle}_{=0} \overset{\varepsilon \to 0}{\longrightarrow} \langle x(t) \rangle = x_0 \tag{2.14}$$

is again constant for all times t. This reminds of Brownian motion, but the constant x_0 cannot be set to zero for GBM, because then the entire process would stay zero. This becomes clear from the recursive definition $x_{j+1} = x_0 \prod_{i=0}^{j} (1 + \varepsilon \sqrt{D} \xi_i)$. The same phenomenon arises in ordinary differential equations of the form $\dot{x}(t) = x(t) f(t)$, where there is $x(0) = 0$. The next step is to derive the standard deviation of the GBM distribution. The expression in Eq. (2.13) results in a simple differential equation for $\langle\langle x(t)^2 \rangle\rangle$. Calculating the

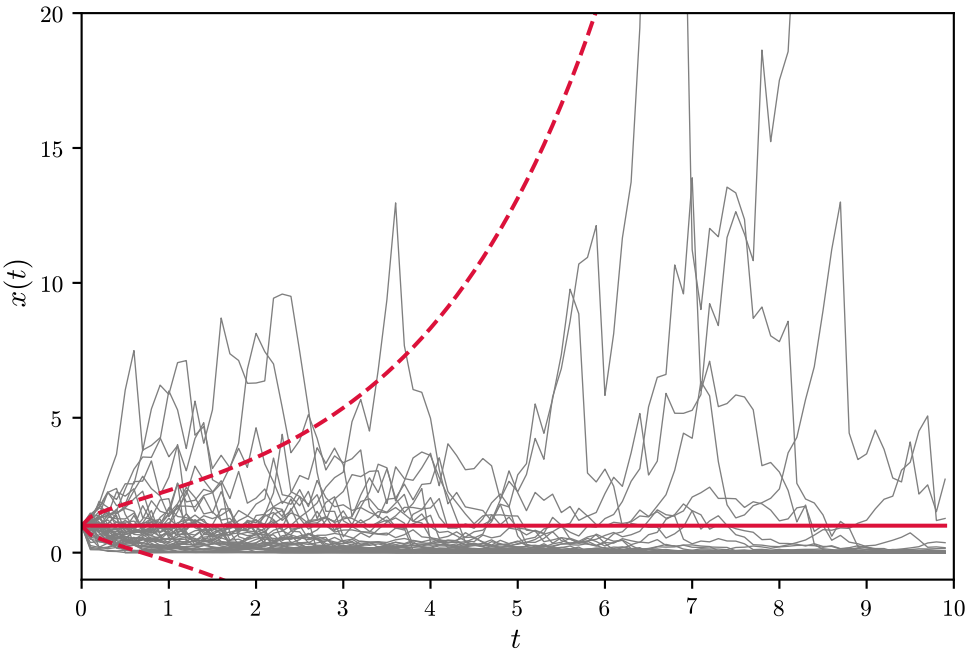


Figure 2.3: Fifty example paths $x(t)$ for discrete geometric Brownian motion over time t in increments of $\varepsilon = 0.1$, starting position $x(0) = 1$, and with a proportionality constant of $D = 1$, within a total time of $T = 10$. The dashed red line indicates the exponential growth of the path distribution's standard deviation $\sigma(t) = \sqrt{e^t - 1}$, while the continuous red line displays the constant mean value at $\langle x(t) \rangle = 1$. This causes a highly asymmetric distribution of the paths, such that most GBM paths land between 0 and 1.

discrete-time variance on both sides leads to

$$\left\langle \left\langle x_{j+1}^2 \right\rangle \right\rangle = \left[1 + \varepsilon^2 D \sigma^2\right] \left\langle \left\langle x_j^2 \right\rangle \right\rangle + \varepsilon^2 D \sigma^2 x_0^2. \tag{2.15}$$

Subtraction of $\langle\langle x_j^2 \rangle\rangle$ on both sides and division by ε delivers a difference equation for the variance, which becomes the ordinary differential equation $\partial \left\langle \left\langle x(t)^2 \right\rangle \right\rangle / \partial t = D \left(\left\langle \left\langle x(t)^2 \right\rangle \right\rangle + x_0^2 \right)$ in the continuum limit. This limit again only properly converges for $\sigma \propto \sqrt{1/\varepsilon}$. The solution is an exponential function. Consequently, the standard deviation is given by

$$\boxed{\sqrt{\langle\langle x(t)^2 \rangle\rangle} = x_0 \sqrt{e^{Dt} - 1}.} \tag{2.16}$$

Figure 2.3 shows a few example paths of GBM asymmetrically distributed around a constant mean value x_0 with exponentially growing variance. Geometric Brownian motion is not normally distributed. Its paths have a constant mean

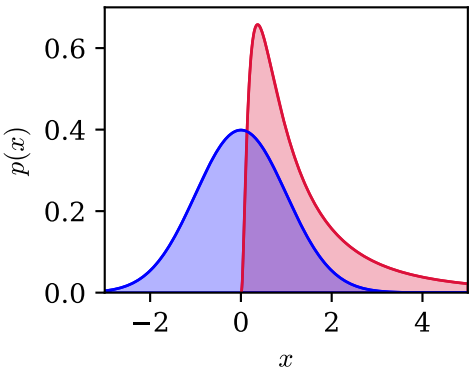


Figure 2.4: Gaussian and log-normal probability density functions. A log-normal distribution of a random variable is defined by the condition that the logarithm of the variable is normally distributed. The mean value and standard deviation of this Gaussian distribution are the characteristic parameters of the log-normal distribution. The figure shows a Gaussian function (blue) with a mean of $\mu = 0$ and a standard deviation of $\sigma = 1$ and the corresponding log-normal probability density function (red) given by $p(x) = \exp\{-[\ln(x) - \mu]^2/2\sigma^2\}/\sqrt{2\pi\sigma^2}x$.

value and a standard deviation growing exponentially in time, while never crossing zero. We need to suspect a highly asymmetric distribution with a large peak between 0 and x_0 and a long tail that causes the large variance. In the analysis of the logarithmic transform of GBM, we can see that it is a *log-normal* distribution. Figure 2.4 shows an example of such a distribution for $x_0 > 0$. The opposite assumption would simply produce an image mirrored at the y-axis.

We set $y_j := \ln(x_j)$. The logarithm turns the product in Eq. (2.13) into a sum producing a recursive definition for the transformed variable sequence $(y_j)_j$ with $y_{j+1} = \ln\left(1 + \varepsilon\sqrt{D}\xi_j\right) + y_j$. This calls for a Taylor expansion of the logarithmic term. For the construction of a SDE for the transformed process, the entire equation is divided by ε later on. Therefore, terms of order $\mathcal{O}(\varepsilon)$ need to be taken into consideration. Furthermore, ξ_j is not a constant in ε. Since it is normally distributed with $\sigma = \sqrt{1/\varepsilon}$, we have $\xi_j \sim \sqrt{1/\varepsilon}$, which leads to

$$\ln\left(1 + \varepsilon\sqrt{D}\xi_j\right) = \underbrace{\varepsilon\sqrt{D}\xi_j}_{\text{stochast.}} - \underbrace{\frac{\varepsilon^2 D\xi_j^2}{2}}_{\text{determ.}} + \mathcal{O}(\varepsilon^{3/2}). \tag{2.17}$$

The first summand is a Wiener process just like for Brownian motion, proving that y_j follows a normally distributed Brownian motion. The second relevant summand does not have stochastic impact. If Eq. (2.17) is inserted back into the recursive relation, it results in the SDE

$$\frac{y_{j+1} - y_j}{\varepsilon} = \sqrt{D}\xi_j - \frac{D}{2}\varepsilon\xi_j^2 + \mathcal{O}(\varepsilon^{1/2}). \tag{2.18}$$

The mean value of the stochastic term vanishes as for Brownian motion in Sec. (2.2). The second summand produces a deterministic drift in the mean position of the transformed path

$$\langle y_{j+1}\rangle - \langle y_j\rangle = \varepsilon\sqrt{D}\underbrace{\langle \xi_j\rangle}_{=0} - \frac{\varepsilon^2}{2}D\underbrace{\langle \xi_j^2\rangle}_{=1/\varepsilon} + \mathcal{O}(\varepsilon^{3/2}). \tag{2.19}$$

Division by ε and application of the continuum limit produces an ordinary first order differential equation with a constant right side. Hence, the mean value of the transformed path $\langle y(t)\rangle = -\frac{D}{2}t$ is linear in t, due to the deterministic force. When it comes to the variance, only the stochastic term has an impact. Since $\langle\langle \xi_j^2\rangle\rangle = 1/\varepsilon$, we obtain the recursive relation $\langle\langle y_{j+1}^2\rangle\rangle = \varepsilon D + \langle\langle y_j^2\rangle\rangle + \mathcal{O}(\varepsilon^{3/2})$ for the variance. In the continuum limit, this yields a variance linear in t and thus a standard deviation of

$$\sqrt{\langle\langle y(t)^2\rangle\rangle} = \sqrt{Dt}, \tag{2.20}$$

exactly as for Brownian motion. The argument of the central limit theorem from Sec. 2.2 can also be applied to this transformed process thus proving that $y(t)$, $t \in \mathbb{R}$, is normally distributed, such that $x(t)$ is log-normally distributed. The logarithm clearly does not commute with taking the mean. The behaviour of the two processes, $(x_j)_j$ and the transform $(y_j)_j$, differs considerably. Figure 2.3 illustrates the behaviour of both processes.

Lastly, we need to make the key observation of this section, the violation of the chain rule. The continuum limit of Eq. (2.18) is an Itō-SDE for $y(t)$ given by

$$\frac{\partial y(t)}{\partial t} = \sqrt{D}\xi(t) - \frac{D}{2}. \tag{2.21}$$

Here, the deterministic term is simply replaced by D as it essentially summarises its effect. The expression in Eq. (2.21) contradicts the chain rule of differential calculus. The chain rule dictates $\partial y(t)/\partial t = \sqrt{D}\xi(t)$. The difference lies in the additional deterministic term that has to be taken into account in Eq. (2.17). It is known as the *Itō term*, because it needs to be considered in Itō SDEs, but not for Stratonovich, as we see in Sec. 2.5.

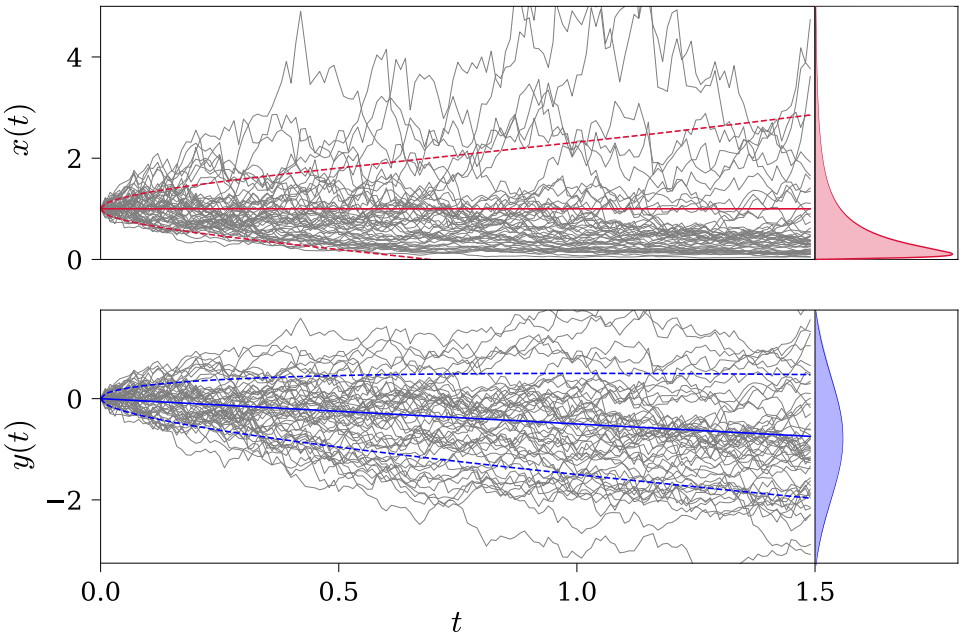


Figure 2.5: Distribution of 50 geometric Brownian motion (GBM) paths $x(t)$ and their logarithmic transform $y(t) := \ln[x(t)]$ after a time of $t = 1.5$ in increments of $\varepsilon = 0.1$ starting at $x(0) = 1$ and $y(0) = 0$. The upper plot displays the behaviour of the GBM paths $x(t)$ with the dotted red line indicating their exponentially growing standard deviation and the continuous red line marking the constant mean at 1. The lower plot shows the logarithm of each of the paths in the upper plot. The continuous blue line in this part indicates the mean value of the transform decreasing linearly with $\langle y(t) \rangle = -t/2$, while the dotted blue line marks the standard deviation around the mean increasing with the square-root of t. On the right, the figure displays the specific distributions of both processes' paths; the log-normal distribution in red on top and the corresponding Gaussian distribution of the transform in blue below. The proportionality constant in the process displayed is $D = 1$.

2.4 General stochastic differential equations

This section is dedicated to examining the effects of varying the random force in x with an arbitrary $\alpha \in [0,1]$ in a general SDE. In this analysis, possible deterministic influences in addition to the stochastic force are taken into consideration. Let $g, f : \mathbb{R}^2 \to \mathbb{R}$ be analytic functions. Then, the SDE

$$\frac{\partial x(t)}{\partial t} = \underbrace{g[x(t),t]}_{\text{determ.}} + \underbrace{f[x(t),t]\xi(t)}_{\text{stoch.}}, \tag{2.22}$$

defines a general one-dimensional stochastic process. The first term g describes an arbitrary deterministic force and f modulates the effect of the white noise on the process. With $\alpha \in [0,1]$, the discrete-time SDE is given by

$$\frac{x_{j+1} - x_j}{\varepsilon} = g(x_j,t) + \left[(1-\alpha)f(x_j,t) + \alpha f(x_{j+1},t)\right]\xi_j. \tag{2.23}$$

For the deterministic force g, the point of evaluation is irrelevant, because g is a simple differentiable function. For a differentiable function, the choice of α is of no importance as discussed in Sec. 2.1. It is the variation in α of the stochastic term and its effect in the continuum limit that is interesting. Itō SDEs can be converted directly into continuous SDEs, as demonstrated in the previous sections. That makes them the preferable choice. An Itō model implies that the incremental stochastic changes of a path depend solely on the position *before* one 'rolls the dice'. However, a system might require a different value of α for its model. This section explores the consequences of allowing the position *after* to have an influence as well. We are going to see that Eq. (2.23) can be reduced to an altered Itō SDE and the choice of α only has deterministic implications in the continuum limit. Proving this claim is the aim of this section.

The positions x_{j+1} and x_j do not differ much, which justifies a Taylor expansion in x_j. The discrete SDE in Eq. (2.23) indicates how much x_{j+1} and x_j differ in orders of ε. For a 'normal' process it would simply be $\Delta x \sim \varepsilon$, but this is incorrect for stochastic processes. The model requires that the variance of the noise $\langle\langle \xi(t)^2 \rangle\rangle$ is antiproportional to ε and thereby $\xi_j \sim \sqrt{1/\varepsilon}$ for all $j \in \mathbb{N}$. As a consequence, the difference between x_{j+1} and x_j is of order $\sqrt{\varepsilon}$ and we need to consider one order more in the expansion than we normally would. That means, the Taylor expansion of f in x_{j+1} at x_j is reduced to

$$f(x_{j+1},t) = f(x_j,t) + f'(x_j,t)[\alpha f(x_{j+1},t) + (1-\alpha)f(x_j,t))]\varepsilon\xi_j + \mathcal{O}(\varepsilon), \tag{2.24}$$

where $f'(x_j,t) := \partial f(x,t)/\partial x|_{x=x_j}$. In Sec. 2.3, it is asserted that $\varepsilon\xi_j^2 = \mathcal{O}(1)$. We can isolate $f(x_{j+1},t)\xi_j$ on the left side and then expand the resulting expression in $\varepsilon\xi_j$. That leaves only two summands in the relevant order, precisely $f(x_{j+1},t)\xi_j = f(x_j,t)\xi_j + f(x_j,t)f'(x_j,t) + \mathcal{O}(\varepsilon^{1/2})$. For the following sections, it is important to recognise that this argument implies $f(x_{j+1},t) = f(x_j,t) + \mathcal{O}(\sqrt{\varepsilon})$ and Eq. (2.23) results in an Itō-SDE

$$\boxed{\frac{x_{j+1}-x_j}{\varepsilon} = g_\alpha(x_j,t) + f(x_j)\xi_j + \mathcal{O}(\varepsilon^{1/2})} \tag{2.25}$$

with an altered deterministic force $g_\alpha(x_j,t) := g(x_j,t) + \alpha f(x_j,t)f'(x_j,t)$. As expected, this yields $g_0(x_j,t) = g(x_j,t)$. This result proves that, in the continuum limit, the stochastic character of a general stochastic process is independent from the choice of α. Consequently, it can be assumed without loss of generality that the process of interest follows an Itō SDE if one is only interested in the stochastic behaviour of the system. Otherwise, the deterministic force needs to be manipulated as shown in this section in order to work with an Itō SDE nevertheless. At last, the choice of α in discrete time has an influence on the behaviour in the form of a deterministic drift, even in the continuum limit. This is a crucial difference to 'normal', solely deterministic processes.

2.5 Variable transformation in stochastic differential equations

The deterministic impact discussed in the previous section is not the only reason information on the α-interpretation is essential when dealing with stochastic processes. At the end of Sec. 2.3, it is pointed out that the logarithm of GBM violates the chain rule of differential calculus. There is an additional term, the *Itō term*. In this final section of the chapter on SDEs, the origins of the Itō term are explored. We are going to see that the Itō term appears in GBM in Sec. 2.3 due to the fact that we have assumed $\alpha = 0$. The outcome is different for the Stratonovich approach of $\alpha = 1/2$. Stratonovich SDEs enjoy significant popularity, because they preserve the chain rule, which makes variable transformations a lot easier to work with. On the other hand, calculating means for and performing the continuum limit on Stratonovich SDEs is significantly more difficult than for Itō. Therefore, the result of the previous section in Eq. (2.25) is an important tool to convert between Itō and Stratonovich depending on the problem.

We take a look at an arbitrary transformation acting on a general stochastic process $\{x(t),\, t \in \mathbb{R}\}$ as defined by Eq. (2.22). The transformation shall be given by a bijective, analytic function $h : \mathbb{R} \to \mathbb{R}$. Then $y(t) := h[x(t)]$ defines the transformed path for $t \in \mathbb{R}$. The goal of this section is to derive a discrete SDE for the transform $(y_j)_{j\in\mathbb{N}}$ of the form

$$\frac{y_{j+1} - y_j}{\varepsilon} = \tilde{g}(y_j, t) + \underbrace{[\alpha \tilde{f}(y_{j+1}.t) + (1-\alpha)\tilde{f}(y_j, t)]}_{=:\Delta\alpha(\tilde{f})}\xi_j + \mathcal{O}(\varepsilon^{1/2}) \tag{2.26}$$

the same α and the same random force that is used for x. This way, the α-interpretation is automatically the same for both paths. In discrete time, we have $y_{j+1} = h(x_{j+1})$ for $j \in \mathbb{N}$, which can be expanded at x_j. Since $(x_{j+1} - x_j) \sim \sqrt{\varepsilon}$, this leads to the discrete SDE

$$\frac{y_{j+1} - y_j}{\varepsilon} = h'(x_j)\frac{x_{j+1} - x_j}{\varepsilon} + \frac{1}{2}h''(x_j)\frac{(x_{j+1} - x_j)^2}{\varepsilon} + \mathcal{O}(\varepsilon^{1/2}), \tag{2.27}$$

with the canonical definitions $h'(x) := \frac{\partial h(x)}{\partial x}$ and $h''(x) := \frac{\partial^2 h(x)}{\partial x^2}$. In order to find $\tilde{g}$ and $\tilde{f}$, both terms on the right side are examined separately. Equation (2.23) implies that only the stochastic term needs to be considered when $(x_{j+1} - x_j)$ is squared since the deterministic term is suppressed with ε. In Sec. 2.4 we find $f(x_{j+1}, t) = f(x_j, t) + \mathcal{O}(\sqrt{\varepsilon})$. Therefore, all that remains of the second term in Eq. (2.27) in the continuum limit is $\frac{1}{2}h''(x_j)f(x_j, t)^2$. That is the *Itō term*. It is solely deterministic, hence part of $\tilde{g}$. The first term in Eq. (2.27) splits into two parts. One is the deterministic term $h'(x_j)g(x_j, t)$. The other one is the stochastic term $h'(x_j)\Delta\alpha(f)\xi_j$ holding all information on α. The stochastic term equals

$$h'(x_j)\Delta\alpha(f)\xi_j = \Delta\alpha(h'f)\xi_j - \alpha[h'(x_{j+1}) - h'(x_j)]f(x_{j+1}, t)\xi_j. \tag{2.28}$$

With the expansion of $h'(x_{j+1})$ in x_j and $\varepsilon\xi_j^2 \equiv 1$ that leaves $h'(x_j)\Delta\alpha(f)\xi_j = \Delta\alpha(h'f)\xi_j - \alpha h''(x_j)f(x_j, t)^2 + \mathcal{O}(\varepsilon^{1/2})$. The first term is the expected stochastic term that can be identified with $\Delta\alpha(\tilde{f})$. The second term is an additional deterministic term that needs to be included in $\tilde{g}$ and which vanishes for $\alpha = 0$. The stochastic process for the transformed path $y(t)$ is consequently defined by the functions

$$\tilde{f}[y(t)] := h'[x(t)]f[x(t), t] \tag{2.29}$$

and

$$\tilde{g}[y(t)] := h'[x(t)]g[x(t),t] + \left(\frac{1}{2} - \alpha\right) h''[x(t)]f^2[x(t),t]. \tag{2.30}$$

From this definition of $\tilde{g}$, it becomes clear why the Itō term appears in Itō SDEs and vanishes for Stratonovich's $\alpha = \frac{1}{2}$. In the continuum limit, the SDE of $y(t)$ is given by

$$\boxed{\frac{\partial y(t)}{\partial t} = \frac{\partial y}{\partial x}\frac{\partial x(t)}{\partial t} + \left(\frac{1}{2} - \alpha\right) h''[x(t)]f^2[x(t),t]} \tag{2.31}$$

if the impact of the stochastic force is varied by α. The chain rule is conserved if and only if we have $\alpha = \frac{1}{2}$ or the transformation fulfils $h''(x) \equiv 0$. We can see an example of the effects of α in GBM as given by Eq. (2.12) with a logarithmic transformation. The corresponding Itō term yields $\frac{1}{2} - \alpha h''[x(t)]f^2[x(t),t] = -\frac{D}{2}$, exactly as calculated in Sec. (2.3).

If the transformation h depends explicitly on the time t, the results of this section still hold, but then the additional term $\dot{h}[x(t),t] := \partial/\partial t\, h[x(t),t]$ needs to be included in Eq. (2.27), Eq. (2.30), and Eq. (2.31) to take the explicit time derivative into consideration.

Chapter 3

Stochastic path integrals

3.1 Construction from probability density functions

In Ch. 2, stochastic processes are modelled by first order differential equations defined by stochastic forces acting as equations of motion. In this chapter, we leave the idea of individual steps controlled by white noise behind. Since the stochastic differential equation (SDE) approach works step by step, it does not immediately provide an understanding of the global character of the system. The path integral approach makes global insights more easily accessible. It no longer focuses on the microscopic dynamic of one step as a small part of a macroscopic path but looks at possible paths as a whole. Path integrals assign a probability to an arbitrary subset of all possible paths of a process. The construction of stochastic path integrals in this section is based on the concepts of the previous section. However, we are going to see that the result is an independent formalism that is equivalent to the SDE approach and produces the same results.

In the previous chapter, a one-dimensional stochastic process is introduced as a family of random variables $(x(t);\ t \geq 0) \subset \mathbb{R}$ in defined by a SDE. The stochastic character of the SDE originates in the white noise random variable $(\xi(t);\ t \geq 0) \subset \mathbb{R}$. Both families have discrete-time representations given by the sequences $(x_j)_{j\in\mathbb{N}} \subset \mathbb{R}$ and $(\xi_j)_{j\in\mathbb{N}_0} \subset \mathbb{R}$ with the discrete-time increment $\varepsilon > 0$. The starting position $x_0 \in \mathbb{R}$ is fixed. The white noise is defined as a family of independent and identically distributed variables with a constant mean of $\langle \xi_j \rangle = 0$ for all $j \in \mathbb{N}_0$ and a standard deviation of $\sqrt{\langle\langle \xi_j^2 \rangle\rangle} = \sqrt{1/\varepsilon}$. Furthermore, we can assume without loss of generality that ξ_j is normally distributed for all $j \in \mathbb{N}_0$ as discussed in Sec 2.2.

We construct the path integrals for a stochastic process in discrete time. The trajectory of a path after a total time $T \in \mathbb{R}$ is divided into $N \in \mathbb{N}$ steps, such that $T = \varepsilon N$. Then, we determine the probability density of the transition $x_j \to x_{j+1}$ for each step $j \leq N$. The product of all these densities results in a density for the entire path, a probability measure over $\mathbb{R}^N$. Figure 3.1 displays an illustration of the procedure. For a fixed total time of $T = \varepsilon N$, a conversion between the N-tuples $\vec{x} := (x_1, \ldots, x_N)$ and $\vec{\xi} := (\xi_0, ..., \xi_{N-1})$ based on the discrete SDE is possible. This can be used to derive the probability density function for $\vec{x}$ from the probability density function of $\vec{\xi}$, which is Gaussian since the white noise is normally distributed. For all $j \in \mathbb{N}_0$, the probability density function of ξ_j is given by the Gaussian function

$$\rho_\xi(\xi_j) := \sqrt{\frac{\varepsilon}{2\pi}} \exp\left(-\frac{\varepsilon}{2}\xi_j^2\right). \tag{3.1}$$

The white noise $(\xi_j)_{j\in\mathbb{N}_0}$ is a sequence of independent and identically distributed variables. Therefore, the probability density function for the vector $\vec{\xi}$ simply equals the product of the probability density functions of all ξ_j, $0 < j < N$, as given by $\rho_{\xi,N}(\vec{\xi}) := \prod_{j=0}^{N-1} \rho_\xi(\xi_j)$. If we define the two vectors of random variables $\vec{\xi} := (\xi_0, ..., \xi_{N-1})$ and $\vec{x} := (x_1, ..., x_N)$ through the Itō SDE

$$\frac{x_{j+1} - x_j}{\varepsilon} = g(x_j, t) + f(x_j, t)\xi_j \tag{3.2}$$

with $t = \varepsilon j$ and analytic functions $g : \mathbb{R}^2 \to \mathbb{R}$ and $f : \mathbb{R}^2 \to \mathbb{R}_{>0}$, this defines a transformation function ϕ mapping between $\vec{\xi}$ and $\vec{x}$, such that $\phi(\vec{x}) = \vec{\xi}$. The SDE in Eq. (3.2) allows us to express ξ_j in terms of x_j and x_{j+1}, which delivers the j-th component of ϕ given by

$$\xi_j = f(x_j, t)^{-1} \left[\frac{x_{j+1} - x_j}{\varepsilon} - g(x_j, t)\right]. \tag{3.3}$$

This leads to a Jacobian matrix $J := (\frac{\partial \xi_i}{\partial x_j})_{ij}$, with $i = 0, ..., N-1$ and $j = 1, ..., N$. The difference in index sets stems from the defining logic of the model, according to which ξ_j and x_{j+1} are correlated, while ξ_j and x_j are not, and the starting point x_0 is not a variable. As a consequence, the matrix entries are given by

$$\frac{\partial \xi_i}{\partial x_j} = \begin{cases} \frac{\partial}{\partial x_i}\left[f(x_i, t)^{-1}\left(\frac{x_{i+1}-x_i}{\varepsilon} - g(x_i, t)\right)\right] & , \text{ if } j = i \geq 1 \\ \frac{\partial}{\partial x_{i+1}}\left[f(x_i, t)^{-1}\left(\frac{x_{i+1}-x_i}{\varepsilon} - g(x_i, t)\right)\right] & , \text{ if } j = i+1 \\ 0 & , \text{else.} \end{cases} \tag{3.4}$$

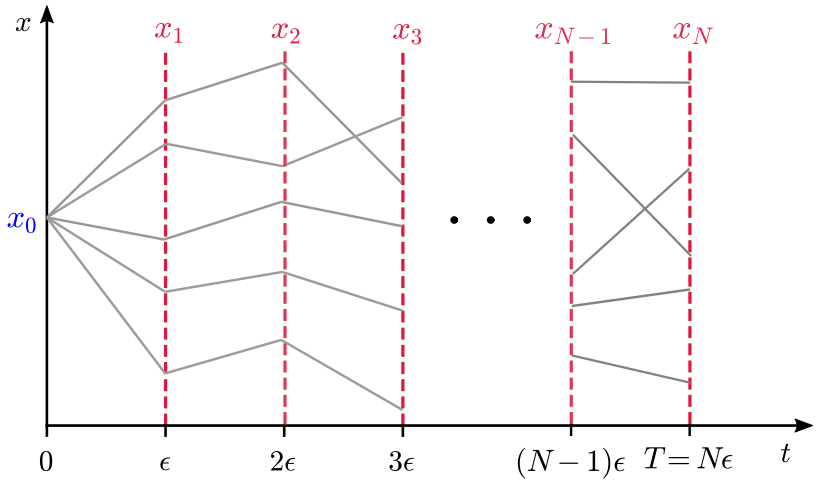


Figure 3.1: Illustration of the construction of a discrete path integral in $N \in \mathbb{N}$ steps. The total time T is divided into N increments $\varepsilon = T/N$. The path starts in position x_0. At each step $j \leq N$, the path integral can take all possible position x_j into account weighting them with their respective transition amplitude. The figure shows five possible example paths.

This means that J is a lower triangular matrix, which simplifies the calculation of the Jacobian determinant $|J|$ significantly. For a triangular matrix, the determinant is given by the product of its diagonal entries. In this particular case, that refers to the product of all $\partial\xi_j/\partial x_{j+1}$ for $j = 0, ...N-1$. Equation (3.3) is linear in x_{j+1} because x_j is considered independent from x_{j+1}. Hence, the Jacobian determinant is given by $|J| = \varepsilon^{-N} \prod_{j=0}^{N-1} 1/f(x_j, t)$. Substitution of $\vec{\xi}$ by $\vec{x}$ in $\rho_{\xi,N}(\vec{\xi})$ and multiplication with $|J|$ yields

$$\rho_{x,N}(\vec{x}) := \exp\left\{-\varepsilon \sum_{j=0}^{N-1} \frac{\left[\frac{x_{j+1}-x_j}{\varepsilon} - g(x_j,t)\right]^2}{2f(x_j,t)^2}\right\} \sqrt{\frac{1}{2\pi\varepsilon}}^N \prod_{j=0}^{N-1} \frac{1}{f(x_j,t)}, \tag{3.5}$$

a normalised probability density function for $\vec{x}$, such that $\rho_{\xi,N}(\vec{\xi})d\vec{\xi} = \rho_{x,N}(\vec{x})d\vec{x}$ is conserved, because we have $d\vec{\xi} = |J|d\vec{x}$. This function acts on $\mathbb{R}^N$ as the product of N Gaussian densities each acting on $\mathbb{R}$. The factors are called *transition amplitudes*. Given x_j, the transition amplitude provides the probability density of the next position x_{j+1}. These densities are Gaussian functions such that

$$\int_{-\infty}^{\infty} \exp\left\{-\varepsilon \frac{\left[\frac{x_{j+1}-x_j}{\varepsilon} - g(x_j,t)\right]^2}{2f(x_j,t)^2}\right\} \sqrt{\frac{1}{2\pi\varepsilon}} \frac{1}{f(x_j,t)} dx_{j+1} = 1 \tag{3.6}$$

holds true for all $j = 0, ..., N-1$. Therefore, the integral over $\rho_{x,N}(\vec{x})$ over all of $\mathbb{R}^N$ also yields 1. Equation (3.5) leads to a *discrete path integral*. If $\mathcal{M} \subset \mathbb{R}^N$ is a subset of all possible N-tuples, the integral over $\rho_{x,N}(\vec{x})$ over $\mathcal{M}$ gives out the probability that an N-step trajectory $\vec{x}$ of the stochastic process starting in x_0 is realised in $\mathcal{M}$. Since it is crucial for the argument in Sec. 3.3, we further need to recognise that Eq. (3.6) implies that the difference $(x_{j+1} - x_j)$ is normally distributed with a standard deviation of $\sqrt{\varepsilon}$, which dictates its typical order of magnitude $(x_{j+1} - x_j) \sim \sqrt{\varepsilon}$.

We can use this discrete path integral to approximate a path integral for a continuous path by performing the continuum limit. Since $T = \varepsilon N$ is fixed, the sum in the exponent multiplied with ε acts like a Riemann sum, which converges to an integral in the continuum limit. In the continuum limit, finite dimensional variable space $\mathbb{R}^N$ approximates infinite dimensional path space $\Omega_x := \lim_{N\to\infty} \mathbb{R}^N$. That requires a new understanding of the differentials in the integral, which leads to the definition of the *path measure*

$$\mathcal{D}[x(t)] := \lim_{\varepsilon\to 0} \sqrt{\frac{1}{2\pi\varepsilon}}^N \, dx_1 \cdots dx_N. \tag{3.7}$$

Holding $T = \varepsilon N$ constant, the continuum limit additionally leads to an infinite product, which is represented by $\mathcal{F}_0\{f[x(t),t]\} := \lim_{\varepsilon\to 0} \prod_{i=0}^{N-1} f(x_i, t)$. The index signals the starting point, which is x_0 in this case. Both structures clearly have the potential to diverge in the continuum limit. The path measure is not well-defined in the continuum limit under any circumstances. Therefore, both $\mathcal{F}_0\{f[x(t),t]\}$ and $\mathcal{D}[x(t)]$ can only be understood symbolically. Nevertheless, their definitions are justified because the integral converges. In the continuum limit, the infinitesimal *path integral* is given by

$$\boxed{\lim_{\varepsilon\to 0} \rho_{x,N}(\vec{x}) d\vec{x} = \exp\left\{ -\int_0^T \frac{\{\dot{x}(t) - g[x(t),t]\}^2}{2f[x(t,t)]^2} dt \right\} \frac{D[x(t)]}{\mathcal{F}_0\{f[x(t),t]\}}.} \tag{3.8}$$

Integration over the space of all possible x-paths Ω_x must equal one because it is the limit of the constant sequence $(\int_{\mathbb{R}^N} \rho_{x,N}(\vec{x}) d\vec{x} = 1)_{N\in\mathbb{N}}$.

The path integrals for Brownian motion and for geometric Brownian motion (GBM) are two prominent examples worth examining for illustration purposes. The stochastic process for Brownian motion as discussed in Sec. 2.2 is defined by $f[x(t),t] \equiv 1$ and $g[x(t),t] \equiv 0$. Inserting this into Eq. (3.8) leads to the

path integral

$$\rho_{x,N}(\vec{x})d\vec{x} \overset{\varepsilon\to 0}{\longrightarrow} \exp\left\{-\int_0^T \frac{\dot{x}(t)^2}{2}dt\right\}\mathcal{D}[x(t)]. \tag{3.9}$$

This exponent ought to remind us of kinetic energy, which hints at a connection to the Feynman path integral formalism of quantum mechanics. If we identify $\xi(t)^2$ with the Lagrangian function $L(q,\dot{q})$ from mechanics, the exponent in Eq. (3.8) can be interpreted as an action integral. Minimal action would thereby imply a maximum of the probability density, which can motivate a stochastic generalisation of Hamilton's principle of least action.

For GBM, we have $f[x(t),t] = x(t)$. In Sec. 2.3, only the Itō case is analysed. However, Sec. 2.4 delivers the procedure for a general $\alpha \in [0,1]$. Therefore, instead of $g[x(t),t] \equiv 0$ as in Sec. 2.3, the deterministic force shall be given as $g[x(t),t] = \alpha f[x(t),t]$ for an arbitrary $\alpha \in [0,1]$. The corresponding path integral for GBM is defined by

$$\rho_{x,N}(\vec{x})d\vec{x} \overset{\varepsilon\to 0}{\longrightarrow} \exp\left\{-\int_0^T \frac{[\dot{x}(t)-\alpha x(t)]^2}{2x(t)^2}dt\right\}\frac{D[x(t)]}{\mathcal{F}_0\{x(t)\}}. \tag{3.10}$$

If the additional terms in the exponent compared to Brownian motion are interpreted as potential energy, this exponent also fits into the action integral scheme of Lagrangian mechanics.

3.2 Hubbard-Stratonovich transformation

The *Hubbard-Stratonovich* (HS) transformation is a practical mathematical tool originally introduced by Ruslan Stratonovich and refined by British physicist John Hubbard. It is an exact transformation based on the principles of completing the square and Gaussian integrals (App. B). Its core appeal lies in the transformation of an expression with a quadratic dependence of a variable μ into one with a linear dependence of that variable. This is achieved by introducing a second independent variable q. The HS transformation is defined by

$$\boxed{\exp\left\{\frac{\mu^2}{2\sigma^2}\right\} = \sqrt{\frac{1}{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left\{-\frac{q^2-2q\mu}{2\sigma^2}\right\}dq.} \tag{3.11}$$

It can be quickly shown that this holds true. Completing the square in the exponent on the right side gives $q^2-2q\mu = (q-\mu)^2+\mu^2$ and the integral over a

Gaussian function over $(-\infty,\infty)$ gives 1, independent of constant shifts to the variable.

Naturally, the HS method can also be applied to multiple variables. Therefore, it can be used for working with path integrals even in the continuum limit. If we have a sequence of variables $(\mu_j)_j$ and the total time $T := \varepsilon N$ for $N \in \mathbb{N}$ steps of $\varepsilon > 0$, then HS yields

$$\exp\left\{\sum_{j=1}^{N} \frac{\mu_j^2}{2\sigma^2}\right\} = \sqrt{\frac{1}{2\pi\sigma^2}}^N \int_{-\infty}^{\infty} \exp\left\{-\sum_{j=1}^{N} \frac{q_j^2 - 2q_j\mu_j}{2\sigma^2}\right\} dq_1 \cdots dq_N. \quad (3.12)$$

If the Gaussian standard deviation for q is chosen to be $\sigma = \frac{1}{\varepsilon}$, the same as for the random force, this integral converges in the continuum limit. With a slightly different symbolic path measure $\widetilde{\mathcal{D}}[q(t)] := \lim_{\varepsilon\to 0} \sqrt{\frac{\varepsilon}{2\pi}}^N dq_1 \cdots dq_N$, the continuum limit of Eq. (3.12) is given by

$$\exp\left\{\int_{-\infty}^{\infty} \frac{\mu(t)^2}{2} dt\right\} = \int_{\Omega_q} \exp\left\{-\int_0^T \frac{q(t)^2 - 2q(t)\mu(t)}{2}\right\} \widetilde{\mathcal{D}}[q(t)], \quad (3.13)$$

where Ω_q represents the space of all q paths that is essentially equivalent to the continuum limit of $\mathbb{R}^N$.

Equation (3.12) can be applied to the probability density in Eq. (3.5) if we chose

$$\mu_j := -i\frac{\frac{x_j - x_{j-1}}{\varepsilon} - g(x_{j-1}, t-\varepsilon)}{f(x_{j-1}, t-\varepsilon)}. \quad (3.14)$$

Making use of the HS transformation simplifies the manipulation of coefficients in the exponent significantly. In the next section, this procedure is crucial to showing that the path integral approach yields the same results for a variable transformation in a stochastic process as the SDE approach from Ch. 2.

3.3 Variable transformation in path integrals

The outcome of Sec. 2.5 makes clear that transformed stochastic processes do not behave according to the rules of differential calculus. This section aims at examining where these deviations come into effect in stochastic path integrals. Let $x(t)$ be the path from Sec. 3.1 whose discrete path integral is defined by the probability density in Eq. (3.5) with the continuum limit in Eq. (3.8). The transformation $h : \mathbb{R} \to \mathbb{R}$ shall be bijective and analytic, such that the transformed path is given by $y(t) := h[x(t)]$ and $x(t) = h^{-1}[y(t)]$ with the inverse

function h^{-1}. Then, Eq. (2.26) provides the SDE for the transformed path. Assuming an Itō interpretation, this leads to the infinitesimal stochastic path integral for y

$$\lim_{\varepsilon\to 0}\rho_{y,N}(\vec{y})d\vec{y} = \exp\left\{-\frac{1}{2}\int_0^T \frac{\{\dot{y}(t)-\tilde{g}[y(t),t]\}^2}{\tilde{f}[y(t),t]^2}dt\right\}\frac{\mathcal{D}[y(t)]}{\mathcal{F}_0\{\tilde{f}[y(t),t]\}}, \tag{3.15}$$

when Eq. (3.8) is applied. The path integral formalism is an alternative approach to stochastic quantisation, equivalent to SDEs. A path integral is essentially a probability density integral. It thereby carries all information on the stochastic behaviour of the system. Therefore, it is possible to produce the path integral for the transformed path $y(t)$ in Eq. (3.15) from the path integral for $x(t)$ through variable transformation, given only the transformation function. That means, it can be shown that Eq. (3.15) equals Eq. (3.8) without insight into the SDEs in particular. Naturally, it is sufficient to show in discrete time that $\lim_{\varepsilon\to 0}[\rho_{x,N}(\vec{x})d\vec{x} - \rho_{y,N}(\vec{y})d\vec{y}] = 0$, which is the goal of this section.

In discrete time, the transformed path is given by $y_j = h(x_j)$. Hence, the discrete differentials of the two sets of variables can be converted into one-another via the corresponding Jacobian such that $d\vec{y} = \prod_{j=1}^N h'(x_j)d\vec{x}$. The change of variables $(x_j)_{j\in\mathbb{N}} \mapsto (y_j)_{j\in\mathbb{N}}$ in the probability density function for x_j in Eq. (3.5) leads to $x_{j+1}-x_j = h^{-1}(y_{j+1})-h^{-1}(y_j)$, which can be expanded in y_j. Equation (3.6) shows that the difference $(x_{j+1}-x_j)$ is normally distributed with a standard deviation of $\mathcal{O}(\sqrt{\varepsilon})$. That implies $(x_{j+1}-x_j)\sim\sqrt{\varepsilon}$. Since, the expansion is divided by ε, all orders up to $\mathcal{O}(\varepsilon)$ need to be taken into account. Accordingly, the expansion in y_j is calculated up to the second order resulting in

$$x_{j+1}-x_j = \frac{1}{h'(x_j)}(y_{j+1}-y_j) - \frac{1}{2}h''(x_j)\frac{(y_{j+1}-y_j)^2}{h'(x_j)^3} + \mathcal{O}(\varepsilon^{3/2}). \tag{3.16}$$

In this expression, we obtain the difference in y that we need for the path integral of the transform $(y_j)_{j\in\mathbb{N}}$ as well as a second term that shows resemblance to the Itō term. In order to prove that the resulting probability density is equivalent to the outcome gained from the construction of $\rho_{y,N}$ from the transformed SDE, we apply the HS transformation to the Gaussian function in the density by identifying μ_j as proposed in Eq. (3.14). The transformation

$$q_j \mapsto \tilde{q}_j := \frac{\tilde{f}(y_{j-1},t-\varepsilon)}{f(x_{j-1},t-\varepsilon)h'(x_j)}q_j \tag{3.17}$$

yields the necessary manipulation of the coefficients, such that $(f,g) \mapsto (\tilde{f},\tilde{g})$. The Jacobi determinant of this transformation cancels the f product and replaces it by the equivalent $\tilde{f}$ product. Furthermore, the h' factor of the Jacobian leads to a cancellation of the Jacobian byproduct that comes with the $(x_j)_{j\in\mathbb{N}} \mapsto (y_j)_{j\in\mathbb{N}}$ conversion. Consequently, the exponential function in Eq. (3.12) is mapped to

$$\exp\left\{-\frac{\varepsilon}{2}\sum_{j=1}^{N}\underbrace{\frac{f(x_{j-1},t-\varepsilon)^2h'(x_j)^2}{\tilde{f}(y_{j-1},t-\varepsilon)^2}}_{=1+\mathcal{O}(\varepsilon^{1/2})}\tilde{q}_j^2\right\}\exp\left\{-\varepsilon\sum_{j=1}^{N}\frac{f(x_{j-1},t-\varepsilon)h'(x_j)}{\tilde{f}(y_{j-1},t-\varepsilon)}\tilde{q}_j\mu_j\right\}. \tag{3.18}$$

Inserting Eq. (3.16) into Eq. (3.14) and multiplying with the transformation coefficient from Eq. (3.17) results in

$$\frac{f(x_{j-1},t-\varepsilon)h'(x_j)}{\tilde{f}(y_{j-1},t-\varepsilon)}\mu_j = -i\frac{\frac{(y_j-y_{j-1})}{\varepsilon}-\tilde{g}(y_{j-1},t-\varepsilon)}{\tilde{f}(y_{j-1},t-\varepsilon)}+\mathcal{O}(\varepsilon^{1/2}). \tag{3.19}$$

The identification $\tilde{g}(y_{j-1}) := g(x_{j-1})h'(x_{j-1}) - \frac{1}{2}h''(x_{j-1})(y_j-y_{j-1})^2/\varepsilon h'(x_{j-1})^2$ leads to the familiar form for a discrete path integral. This identification is justified, because with the results from Sec. 2.5 it can quickly be proven that the second term is really equivalent to the Itō term. We get

$$\frac{1}{2}h''(x_j)\frac{(y_{j+1}-y_j)^2}{\varepsilon h'(x_j)^2} = \frac{1}{2}h''[x(t)]f^2[x(t),t]+\mathcal{O}(\varepsilon^{1/2}). \tag{3.20}$$

We can summarise that variable transformation in the stochastic path integral formalism yields the same results as the SDE approach. While SDEs offer valuable insight into the origin of the dynamic of a process, the path integral approach works with simple numbers and does not require knowledge of the white noise $\{\xi(t), t\in\mathbb{R}\}$.

Chapter 4

Fokker-Planck equation

4.1 Derivation from path integrals

This chapter introduces a third, more abstract approach to modelling stochastic processes. While stochastic differential equations describe the microscopic dynamic of a stochastic process and path integrals assign a probability to specific trajectories, the Fokker-Planck approach only examines the probability of the end position x of an arbitrary path after a time t. The *Fokker-Planck equation* (FPE) is a second order partial differential equation of the probability density $p(x,t)$ of a path being in position x at time t. It is named after the physicists Adriaan Daniël Fokker and Max Planck. We are deriving the FPE using the results of the previous chapters, but the result is going to be an independent formalism equivalent to the other two. Given a stochastic process as defined by the SDE in Eq. (3.2), the goal of this section is to reproduce the corresponding FPE defined by

$$\frac{\partial}{\partial t}p(x,t) = -\frac{\partial}{\partial x}\left[p(x,t)g(x,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[p(x,t)f(x,t)^2\right]. \tag{4.1}$$

The starting point x_0 is considered a presupposition of the system. The probability density $p(x,t)$ can be calculated by expressing the probability density of an arbitrary path in discrete time as given by Eq. (3.5) and then integrating over all steps but the last. Only the end point x at t is of interest not the arbitrary path which leads there. Figure 4.1 offers a visualisation of this set up.

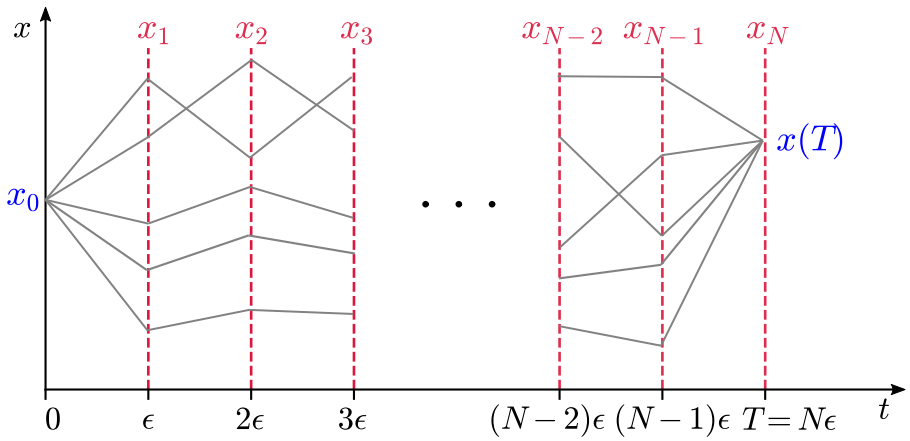


Figure 4.1: Illustration of a discrete path integral over all but the last step of an N-step trajectory with $N \in \mathbb{N}$ and a time increment of $\varepsilon = T/N$ for a fixed time $T > 0$. The t-axis described the progress in time while the x-axis displays the position. The path starts at x_0 and ends in $x(T)$. Each of the $(N-1)$ steps x_j, $j < N$, in between is integrated over all of $\mathbb{R}$ and weighted with its respective transition amplitude. The figure shows five possible example paths.

With the probability density function for the entire path defined in Eq. (3.5), the probability density for arriving in x_{j+1} at t_{j+1} for $j \in \mathbb{N}$ is given by

$$p(x_{j+1}, t_{j+1}) := \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{j} \underbrace{\rho_{x,j+1}(x_1, ..., x_{j+1})}_{=:\rho_{x,j}(x_1,...,x_j)\Psi(x_j, x_{j+1}-x_j)} \, dx_1 \cdots dx_j, \tag{4.2}$$

where $\Psi(x, t, \Delta x)$ represents the probability density for the transition $x \to x + \Delta x$, the transition amplitude. As discussed in Sec. (3.1), the density $\rho_{x,j+1}(x_1, ..., x_{j+1})$ is made up of a product of one-dimensional Gaussian distributions for each step of the path. The product of all but the last step's transition amplitudes forms the probability density for arrival in x_j at t_j, which leads to the recursive expression $p(x_{j+1}, t_{j+1}) = \int_{-\infty}^{\infty} p(x_j, t_j)\Psi(x_j, t_j, x_{j+1} - x_j)dx_j$ with the transition amplitude

$$\Psi(x, t, \Delta x) := \exp\left\{ -\frac{\varepsilon}{2} \frac{\left[\frac{\Delta x}{\varepsilon} - g(x,t)\right]^2}{f(x,t)^2} \right\} \frac{\sqrt{\frac{1}{2\pi\varepsilon}}}{f(x,t)}. \tag{4.3}$$

This expression is the key argument of the derivation in this section. A change of variables by $x_j \mapsto \Delta x := x_{j+1} - x_j$ motivates a Taylor expansion of the integrand $p(x_{j+1} - \Delta x, t_{j+1} - \varepsilon)\Psi(x_{j+1} - \Delta x, t_{j+1} - \varepsilon, \Delta x)$ in $(x_j, t_j) \approx (x_{j+1}, t_{j+1})$. In the transition amplitude $\Psi(x, t, \Delta x)$, x and Δx need to be considered as

independent variables, such that the expansion in x_{j+1} only affects the first, because the integral over $\Psi(x_{j+1}, t_{j+1}, 0)$ does not converge. Also, it intuitively makes sense to consider the position and the transition interval as independent from each other. From the Gaussian form of the distributions, it has been found that $\Delta x \sim \sqrt{\varepsilon}$. Therefore, the expansion results in

$$\begin{aligned} p(x_{j+1}, t_{j+1}) =& p(x_{j+1}, t_{j+1}) \int_{-\infty}^{\infty} \Psi(x_{j+1}, t_{j+1}, \Delta x)\, d\Delta x \\ &- \frac{\partial}{\partial x}\left[p(x,t) \int_{-\infty}^{\infty} \Delta x \Psi(x, t, \Delta x)\, d\Delta x\right]_{(x_{j+1}, t_{j+1})} \\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[p(x,t) \int_{-\infty}^{\infty} (\Delta x)^2 \Psi(x, t, \Delta x)\, d\Delta x\right]_{(x_{j+1}, t_{j+1})} \\ &- \varepsilon \frac{\partial}{\partial t}\left[p(x,t) \int_{-\infty}^{\infty} \Psi(x, t, \Delta x)\, d\Delta x\right]_{(x_{j+1}, t_{j+1})} + \mathcal{O}(\varepsilon^{3/2}). \end{aligned} \tag{4.4}$$

The zeroth order cancels with the left side, as the integral equals one because $\Psi(x, \Delta)$ is a Gaussian function in Δx. For the same reason, all that remains of the first order term in t is the derivative of p. The integral factors in the first and second order terms in x equal the mean values $\langle \Delta x \rangle_\Psi$ and $\langle (\Delta x)^2 \rangle_\Psi$ under the Ψ-distribution. Since $\Psi(x, t, \Delta x)$ is a Gaussian function in Δx, we can conclude

$$\langle \Delta x \rangle_\Psi = \varepsilon g(x, t) \tag{4.5}$$

from the mean given by the exponent and

$$\langle (\Delta x)^2 \rangle_\Psi = \varepsilon f(x, t)^2 + \varepsilon^2 g(x, t)^2. \tag{4.6}$$

from the standard deviation. Division by ε and the performance of the continuum limit result in the FPE in Eq. (4.1).

This representation of the stochastic process displays a striking resemblance to the Schrödinger equation. We can define a 'Hamiltonian' F as the operator

$$F\left(\frac{\partial^2}{\partial x^2}, \frac{\partial}{\partial x}, x, t\right) := -\left(\frac{\partial}{\partial x}\right)[g(x,t)] + \frac{1}{2}\left(\frac{\partial^2}{\partial x^2}\right)[f(x,t)^2]. \tag{4.7}$$

This way, Eq. (4.1) turns into $\frac{\partial}{\partial t} p(x,t) = F p(x,t)$, a Schrödinger equation in all but the missing imaginary unit i on the left side. As such a second order partial differential equation, the FPE is fairly easy to work with and to solve via separation of variables and superposition, which makes it an attractive approach to stochastic quantisation. If g and f are allowed complex values,

the FPE can be understood as a more general equation that also covers the Schrödinger equation. The Fokker-Planck operator in Eq. (4.7) however is not generally Hermitian. This can be quickly seen for the simple, non-stochastic case of $g(x,t) \equiv \text{const.}$ and $f(x,t) \equiv 0$.

Diffusive systems are an important application of stochastic processes. The Fokker-Planck model is pivotal for thermodynamics. For $g(x,t) \equiv 0$ and $f(x,t)^2 \equiv D \in \mathbb{R}_{\neq 0}$, the FPE equals the diffusion equation. For this reason, the function $f(x,t)^2$ is also called the *diffusion coefficient* and $g(x,t)$ the *drift coefficient*. In both contexts, quantum mechanics and thermodynamics, it is further appropriate to interpret the FPE as a continuity equation of the form

$$\frac{\partial}{\partial t}\rho = -\frac{\partial}{\partial x}j \tag{4.8}$$

with a density of $\rho := p(x,t)$ and a flux of $j := [p(x,t)g(x,t)] - \frac{1}{2}\frac{\partial}{\partial x}[p(x,t)f(x,t)^2]$.

Brownian motion and geometric Brownian motion both describe diffusive systems. In correspondence with the model presented in Sec. 2.2, Eq. (4.1) yields the FPE for Brownian motion given by

$$\frac{\partial}{\partial t}p(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2}p(x,t) \tag{4.9}$$

which is solved by the Gaussian density $p(x,t) = \sqrt{\frac{1}{2\pi Dt}}\exp\left\{-\frac{x^2}{2Dt}\right\}$, the same result as found in Sec. 2.2. For GBM, the model in Sec. 2.3 implies the FPE

$$\frac{\partial}{\partial t}p(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[p(x,t)x(t)^2\right] \tag{4.10}$$

which is solved by the log-normal density $p(x,t) = \frac{1}{x}\sqrt{\frac{1}{2\pi Dt}}\exp\left\{-\frac{[\ln(x)+Dt/2]^2}{2Dt}\right\}$, the same result as found in Sec. 2.3. These two examples support the fact that the FPE approach is equivalent to SDEs and path integrals.

4.2 Variable transformation in Fokker-Plank equations

As for the previous two approaches to stochastic quantisation, we want to examine how the FPE behaves under variable transformation in order to show that the FPE approach independently produces the same results as SDEs and path integrals. In this formalism, a variable transformation of the process is

going to be significantly easier to perform than in the other two approaches. We are only going to need basic calculus. All the calculations of this section are going to be strictly analytical.

The transformed variable shall again be given by $y(t) := h[x(t)]$ for a bijective analytic transformation $h : \mathbb{R} \to \mathbb{R}$ and a stochastic process $(x(t),\ t \geq 0)$ defined by the SDE in Eq. (3.2). Based on the results from Sec. 2.5, this leads to the FPE for $y(t)$ given by

$$\frac{\partial}{\partial t}\tilde{p}(y,t) = -\frac{\partial}{\partial y}\left[\tilde{p}(y,t)\tilde{g}(y,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[\tilde{p}(y,t)\tilde{f}(y,t)^2\right], \tag{4.11}$$

where we have the drift and diffusion coefficients from Eq. (2.29) and Eq. (2.30). It is possible the produce this same equation for the transformed process directly from the FPE for $x(t)$ without prior knowledge of neither the corresponding SDE nor the stochastic path integrals. We begin the proof by substituting $x = h^{-1}(y)$ with the inverse transformation function h^{-1} in Eq. (4.1). This leads to

$$\begin{aligned}\frac{\partial}{\partial t}p[h^{-1}(y),t] =& -h'[h^{-1}(y)]\frac{\partial}{\partial y}\left\{g[h^{-1}(y)]p[h^{-1}(y),t]\right\} \\ &+\frac{1}{2}h''[h^{-1}(y)]\frac{\partial}{\partial y}\left\{f[h^{-1}(y),t]^2p[h^{-1}(y),t]\right\} \\ &+\frac{1}{2}h'[h^{-1}(y)]^2\frac{\partial^2}{\partial y^2}\left\{f[h^{-1}(y),t]^2p[h^{-1}(y),t]\right\},\end{aligned} \tag{4.12}$$

where the primed notation indicates derivatives after x, that is $h'(x) = \frac{\partial h(x)}{\partial x}$ and $h''(x) = \frac{\partial^2 h(x)}{\partial x^2}$. Since the variables are equivalent, the original probability density function $p(x,t)$ needs to be equivalent to the one of the transformed process $\tilde{p}(y,t)$. Via substitution one can be converted into the other. For all integration sets, the identity $\tilde{p}(y,t)dy = p(x,t)dx$ must be retained. Since the transformation function implies $\frac{\partial y}{\partial x} = h'(x)$, the probability density function of the transformed process is given by $\tilde{p}(y,t) = p[h^{-1}(y),t]/h'[h^{-1}(y)]$. Implementing this result into Eq. (4.12) leads to

$$\begin{aligned}\frac{\partial}{\partial t}\tilde{p}(y,t) =& -\frac{\partial}{\partial y}\left\{h'[h^{-1}(y)]g[h^{-1}(y),t]\tilde{p}(y,t)\right\} \\ &+\frac{1}{2}\frac{h''[h^{-1}(y)]}{h'[h^{-1}(y)]}\frac{\partial}{\partial y}\left\{h'[h^{-1}(y)]f[h^{-1}(y),t]^2\tilde{p}(y,t)\right\} \\ &+\frac{1}{2}h'[h^{-1}(y)]\frac{\partial^2}{\partial y^2}\left\{h'[h^{-1}(y)]f[h^{-1}(y),t]^2\tilde{p}(y,t)\right\}.\end{aligned} \tag{4.13}$$

With $\frac{\partial}{\partial y}h'(x) = h''(x)/h'(x)$, the product rule, and identifying $\tilde{f}$ according to Eq. (2.29), we find that the final term in Eq. (4.13) equals the second term in Eq. (4.11) plus additional terms from the product rule. The remainder cancels with the second term in Eq. (4.13) entirely but for the Itō term, which combines with the first term to form $\tilde{g}$ corresponding to Eq. (2.30). The product rule yields

$$\begin{aligned} &\frac{1}{2}h'(x)\frac{\partial^2}{\partial y^2}\left\{\frac{\tilde{f}(y,t)^2}{h'(x)}\tilde{p}(y,t)\right\} + \frac{\partial}{\partial y}h'(x)\frac{\partial}{\partial y}\left\{\frac{\tilde{f}(y,t)^2}{h'(x)}\tilde{p}(y,t)\right\} \\ =&\frac{1}{2}\frac{\partial^2}{\partial y^2}\left\{\tilde{f}(y,t)^2\tilde{p}(y,t)\right\} - \frac{1}{2}\frac{\tilde{f}(y,t)^2}{h'(x)}\tilde{p}(y,t)\frac{\partial^2}{\partial y^2}h'(x). \end{aligned} \tag{4.14}$$

For the sake of the equation's clarity, the variable x is not substituted by $h^{-1}(y)$ in some terms. The reoccurring term in Eq. (4.13) and in Eq. (4.14) produces the Itō term and cancels the remainder on the right side of Eq. (4.14) by

$$\begin{aligned} &-\frac{1}{2}\frac{h''(x)}{h'(x)}\frac{\partial}{\partial y}\left\{\frac{\tilde{f}(y,t)^2}{h'(x)}\tilde{p}(y,t)\right\} \\ =&-\frac{1}{2}\frac{\partial}{\partial y}\left\{h''(x)f(h^{-1}(y),t)^2\tilde{p}(y,t)\right\} + \frac{1}{2}\frac{\tilde{f}(y,t)^2}{h'(x)}\tilde{p}(y,t)\frac{\partial^2}{\partial y^2}h'(x). \end{aligned} \tag{4.15}$$

That leaves us with the FPE for the transformed path $y(t)$ as given in Eq. (4.11). Performing a variable transformation for FPEs only requires differential calculus. It can be performed in continuous time without consideration of incremental steps $\varepsilon > 0$. We can conclude that, by parting from the tangible notion of trajectories, the advanced level of abstractness of the FPE model creates a structure which is complete and analytically sound.

Chapter 5

Conclusion and outlook

5.1 Summary

In this thesis, three different approaches to modelling stochastic processes as the basis of stochastic quantisation have been explored. Chapter 2 has introduced the concept of a stochastic differential equation (SDE) which describes the microscopic dynamic of a stochastic process. In Ch. 3, we have constructed path integrals in order to measure the probability of specific paths. The Fokker-Planck approach in Ch. 4 has reduced the stochastic process to the probability of a specific endpoint of an arbitrary path of the process. This section summarises the results of the three approaches obtained throughout this thesis. Table 5.1 displays an overview of the key points.

The SDE approach has caused mathematical complications in continuous time, because stochastic processes are mostly not differentiable. This issue has motivated the quantisation of time, since difference equations are well-defined nevertheless. An SDE is characterised by a random force, the cause of the non-differentiability. The random force has been established as a function of the random variable $\xi(t)$ known as white noise. We have studied the characteristics of white noise as the origin of the stochastic nature of an SDE throughout Ch. 2. Most importantly, we have discovered that the variance of the noise variable in discrete time needs to be scaled anti-proportional to the time increment ε, with which time is discretised, in order for the model to hold in the continuum limit $\varepsilon \to 0$. Meanwhile, two important examples of stochastic processes have been presented, Brownian motion and geometric Brownian motion. We have found that the paths of Brownian motion are normally distributed with a constant mean value and a standard deviation growing with the square-root of time. Geometric Brownian motion has proven to follow a log-normal distribution and

we have seen its logarithmic transform can be described as a Brownian motion process with a deterministic drift proportional to the time passed.

Discretising time has been at the centre of most calculations in this thesis. It has become clear in Ch. 2 that, for stochastic processes, the discrete-time representation of their corresponding SDE is not unique, because the processes are not differentiable. As a consequence, we have learned that converting SDEs to discrete time requires additional information, that is on how to interpret $\alpha \in [0, 1]$ in a discrete-time SDE such as

$$\frac{x_{j+1} - x_j}{\varepsilon} = g(x_j, t) + [(1 - \alpha) f(x_j, t) + \alpha f(x_{j+1}, t)]\, \xi_j \tag{5.1}$$

discussed in Sec. 2.4. The analysis in Sec. 2.4 has concluded that the particular choice of α for the model of a stochastic process has deterministic impact only. Therefore, all SDEs can be converted into an Itō SDE, that is assuming $\alpha = 0$, which is the most convenient choice as to making predictions based on the SDE of a process alone. However, in the case of a variable transformation, choosing an Itō SDE results in a violation of the chain rule of differential calculus which is only preserved for Stratonovich SDEs, that is for $\alpha = \frac{1}{2}$, as Sec. 2.5 has shown.

In Ch. 3, we have developed an approach to stochastic quantisation which adopts the paths of a stochastic process as a whole instead of a step by step analysis as attempted by SDEs. Based on the Gaussian probability density function of the white noise variable $\xi(t)$, we have constructed path integrals producing the probability of a random path being in a specific subset of all possible paths. Discretising time has been the foundation of this construction. Starting from a finite product of Gaussian-like functions as transition amplitudes for each step, the continuum limit has yielded an approximation of a continuous path integral over an approximated path space. In this formalism, the white noise and its stochastic nature are only implied through the character of the integrals, but not explicitly included in the theory. This simplifies the theory, since no further knowledge of the process' stochastic behaviour is required for its calculation, only simple numeric calculus. Furthermore, the integral approach makes global insights into the process more easily accessible than SDEs since it identifies whole paths as single events.

Lastly, the Fokker-Planck equation (FPE) introduced in Ch. 4 has offered an approach to stochastic quantisation which reduces the entire dynamic of a stochastic process to a single, continuous partial differential equation, a likeness of the

	SDE	**path integral**	**Fokker-Planck**
concept	- microscopic dynamic - white noise	- probability of entire paths	- no paths - probability of endpoints
strengths	- ideal for simulations - quick calculation of mean and standard deviation	- simple calculations - global insight	- non-stochastic equation - analytic calculations
weaknesses	- lack of global insights - unknown distribution - semiclassical assumptions	- possibly costly calculations - semiclassical assumptions	- no insights into classical trajectories

Table 5.1: Overview of the three approaches to stochastic quantisation including their respective strengths and weaknesses. The first column summarises the results obtained from stochastic differential equations (SDE) approach, the second focuses on stochastic path integrals and the third on Fokker-Planck equations.

Schrödinger equation. This second order partial differential equation has proven non-stochastic in and of itself. It is the object of the differential equation which has been identified with the probability density function $p(x,t)$ of a path arriving in position x at time t. The path-integral approach has provided a practical expression of $p(x,t)$ by integrating over all possible paths beginning in x_0 and ending in x. This expression has made the derivation of the FPE possible. Nevertheless, the resulting formalism for modelling stochastic processes based on FPEs, is independent from and equivalent to stochastic path integrals as well as SDEs. In the Fokker-Planck formalism, variable transformations have proven to be particularly easy. A general variable transformation has been performed using only basic analytical calculations, which has yielded the same result as the more costly calculations in the other two formalisms. Since, the FPE approach dismisses concrete paths, it elevates to a level of abstractness which allows to model processes for which there are no actual paths, as we see it for example in quantum mechanics.

All three approaches have their particular merits and downfalls, much like the Lagrangian, the Hamilton and the Hamilton-Jacobi formalisms do for mechanics or Feynman's path integral approach and Dirac's operator approach for quantum mechanics. Having multiple equivalent options to chose from has the benefit of being able to adjust the approach depending on the problem.

5.2 Parallelism to quantum mechanics

We have come across many indicators of parallel structures behind quantum mechanical systems and stochastic processes throughout this thesis. For example, we have seen in Sec. 2.2 that measurements of the position and the velocity of Brownian motion are noncommutative, a defining phenomenon between the position and momentum operator in the canonical quantisation of quantum mechanics. Furthermore, the specific outcome of a stochastic process is unpredictable just like the outcome of a double-split experiment and Ch. 4 has even produced a version of the Schrödinger equation for transition amplitudes of stochastic processes. Attempting stochastic quantisation at least as part of a semiclassical quantum theory seems appropriate. After all, quantum mechanics, as the name indicates, has a long history of quantisation approaches. After Planck and Einstein postulated discrete energy quants in radiation fields, Bohr went on basing his atomic model on quantised angular momentum $L_n = \hbar n$, $n \in \mathbb{N}$. Sommerfeld streamlined this idea in 1915 and demanded a quantisation of the action integral of every degree of freedom i by $S_i = 2\pi\hbar n_i$ [8]. Finally, Paul Dirac motivated the pathintegral formulation of quantum mechanics, a similar approach to stochastic quantisation, which was later fully developed by Richard Feynman.

The canonical quantisation of quantum mechanics is based on states represented by elements of a Hilbert space. It usually assumes that a specific state is prepared and then aims at determining the probability with which a measurement produces a certain outcome. A state is not a path or a set of other specific observable quantities. A quantum state is an abstract object which implicitly holds information on the probability of outcomes of measurements. Possible paths are not seen as individual trajectories but as a spectrum of possibilities, some subset of all of path space mathematically speaking, but paths are not actually realised. Therefore, the classical quantisation, the centre of which is the Schrödinger equation, can be associated with the Fokker-Planck approach introduced in Ch. 4.

The path integral formulation of quantum mechanics is based on generalising the action principle of Lagrangian mechanics and it works very similar to the stochastic path integral approach in Ch. 3. For the construction of the path integral quantisation of quantum mechanics, infinitesimal transitions are used to construct a macroscopic dynamic. Like in Ch. 2, a fixed total time $T \in \mathbb{R}$ is divided into a finite amount of small increments $\varepsilon = T/N$, $N \in \mathbb{N}$, which is later

used to perform a continuum limit, meaning $\varepsilon \to 0$ as well as $N \to \infty$ given that T is fixed. Accordingly, a path from x_0 to x_N is split up into $N >> 1$ short, linear paths. Then, the transition amplitude for the finite time T can be approximated by a product of the amplitudes of an infinite amount of infinitesimal steps. Assuming a Hamiltonian of $H = \frac{p^2}{2m} + V(x)$, where p is the momentum and x is the position operator and V a potential, the transition amplitude for an infinitesimal step $\varepsilon > 0$ yields

$$\langle x_j | e^{-iH\varepsilon} | x_{j+1} \rangle = \frac{1}{\sqrt{2\pi i\varepsilon}} \exp\{i\Delta S_j\},\ j \in \mathbb{N}, \tag{5.2}$$

with an infinitesimal action integral

$$\Delta S_j = \varepsilon \left[\frac{m}{2} \left(\frac{x_j - x_{j-1}}{\varepsilon} \right)^2 - V\left(\frac{x_j + x_{j-1}}{2} \right)^2 \right], \tag{5.3}$$

which resembles a Lagrangian, resulting in a total transition amplitude of

$$\langle x_0 | e^{-iHT} | x_N \rangle = \lim_{\varepsilon \to 0} \left(\frac{1}{\sqrt{2\pi i\varepsilon}} \right)^{N/2} \int \exp\left\{ i \sum_{j=1}^{N} \Delta S_j \right\} \prod_{j=1}^{N-1} dx_j \tag{5.4}$$

for all possible paths allowing $(x_0, 0) \mapsto (x_N, T)$ [9]. This construction follows the same notion as the path integral formulation for stochastic processes in Ch. 3. It bares similar pitfalls as the path integrals in Sec. 3.1 as well. We need to stress that the performance of the continuum limit on a structure like Eq. (5.4) can only by understood symbolically since the limit is not necessarily well-defined. The 'velocity' $v = \Delta x/\varepsilon$ in $\sum_j \Delta S_j \to S[x(t)] = \int_0^T dt[\frac{m}{2}v^2 - V(x)]$ possibly diverges. Additionally, the integration measure is not technically defined for an infinite amount of steps just like in Sec. 3.1. Nevertheless, the transition amplitude holds for the continuum limit and is thereby deemed physically meaningful [9]. The path integral quantisation enables approximative calculations necessary in the realm of quantum mechanics and quantum field theory that would be impossible to solve in the abstract operator formalism of the canonical quantisation. However, it is a semiclassical model based on the premise of the existence of actual paths, which is not necessarily given.

Appendix A

Stochastic processes in probability theory

A stochastic process defines a family of random variables in a probability space. This chapter briefly summarises the mathematical structures needed to construct a stochastic process.

A **probability space** is defined as a triple $(\Omega, \mathcal{A}, P)$ consisting of a *sample space* $\Omega \neq \emptyset$ holding all possible outcomes, a σ-algebra $\mathcal{A} \subset \mathcal{P}(\Omega)$ called *event space*, where $\mathcal{P}(\Omega)$ is the power set of Ω, and a *probability function* $P : \mathcal{A} \to [0, 1]$. The probability function P must satisfy the Kolmogorov axioms

(i) $P(A) \geq 0$ for all $A \in \mathcal{A}$,

(ii) $P(\Omega) = 1$ and

(iii) $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$ for all mutually disjoint sets $(A_n)_n \subset \mathcal{A}$.

The pair made of $(\Omega, \mathcal{A})$ is called *measurable space* [10].

Given a probability space $(\Omega, \mathcal{A}, P)$ and a measurable space $(\Omega', \mathcal{A}')$, a **random variable** is defined as a $\mathcal{A}$-$\mathcal{A}'$-measurable map $X : \Omega \to \Omega'$. The measurability is given if $X^{-1}(A') \in \mathcal{A}$ for all $A' \in \mathcal{A}'$ [10].

A stochastic process is a set of random numbers, but not necessarily a sequence. Its index set does not need to be countable. This is important, since stochastic processes are used to model events with a continuous development in time. Therefore, a general **stochastic process** is defined as a quadruple $(\Omega, \mathcal{A}, P, (X_t)_{t \in I})$, where $(\Omega, \mathcal{A}, P)$ is a probability space and $(X_t)_{t \in I}$ is a family of random variables acting on this probability space with values in a common

measurable space $(E, \mathcal{B})$. The index set $I \neq \emptyset$ is called the *parameter set* and the measurable space $(E, \mathcal{B})$ is called *state space*. For all $\omega \in \Omega$ the map $I \to E, t \mapsto X_t(\omega)$ defines a *path* of the process. Typically, the state space E is a subspace of $\mathbb{R}^N$ and the index set equals the set of positive real numbers $I = \mathbb{R}_+$, then often referred to as the *time set*. Paths of a process are also called *trajectories* or *realisations* of the process [11].

A **Wiener process** is a special continuous-time stochastic process $(X_t)_{t\geq 0}$ characterised by a constant mean value of 0, independent and identically distributed Gaussian increments and continuous paths. In a continuous-time process, an increment is defined as a difference $(X_t - X_{t'})$ for $0 \leq t < t'$ [12].

Appendix B

Gaussian integrals in $\mathbb{R}^n$

This chapter offers a superficial review of Gaussian integrals. The simplest form of a Gaussian function is given by $f(x) := \exp\big(-x^2\big)$. The corresponding Gaussian integral over $(-\infty, \infty)$ yields $\sqrt{\pi}$. Substitution by $x \mapsto \sqrt{\frac{1}{2}a}x$ leads to the typical Gaussian 'bell curve'

$$I := \int_{-\infty}^{\infty} \exp\Big\{-\frac{1}{2}ax^2\Big\}\, dx = \sqrt{\frac{2\pi}{a}}. \tag{B.1}$$

This result is invariant under linear shifts such as $x \mapsto (x - \mu)$ for $\mu \in \mathcal{R}$, since this transformation produces a Jacobian of 1. This knowledge of simple Gaussian functions can be used to solve similar integrals with arbitrary polynomials of the second order in the exponent. The integral shall be given by

$$I_1 := \int_{-\infty}^{\infty} \exp\big\{-\alpha x^2 - \lambda x + c\big\}\, dx \tag{B.2}$$

with parameters $\alpha \in \mathbb{R}_+$, $c \in \mathbb{R}$ and $\lambda \in \mathbb{R} \setminus \{0\}$. Completing the square yields $x^2 - \frac{\lambda}{\alpha}x = (x - \frac{\lambda}{2\alpha})^2 - \frac{\lambda^2}{4\alpha^2}$ which results in $I_1 = \exp\Big(c + \frac{\lambda^2}{4\alpha}\Big)\sqrt{\frac{\pi}{\alpha}}$.

Similar arguments can be applied to higher dimensions $n > 1$. We look at an integral

$$I_n := \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\Big\{-\vec{x}^T M \vec{x} + \vec{\lambda}^T \vec{x} + c\Big\}\, dx_1 \cdots dx_n \tag{B.3}$$

with $\vec{x} := (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, $\vec{\lambda} \in \mathbb{R}^n$, $c \in \mathbb{R}$ and an invertible, diagonalisable matrix $M \in \mathbb{R}^{n \times n}$, which can be chosen symmetric without loss of generality for any other part would vanish in $-\vec{x}^T M \vec{x}$.

First, we consider the case of $\vec{\lambda} \equiv 0$. Since M is diagonalisable, there is an base-change matrix $S \in \mathbb{R}^{n\times n}$, such that $D := SMS^T$ is diagonal with entries $d_i := D_{ii}$, $1 \leq i \leq n$. S is orthogonal in that $S^T = S^{-1}$. We define the transformed variable $\vec{y} := S\vec{x}$. Then, a substitution in the integral leads to a familiar Gaussian function. The Jacobian for the transformation is given by $|\det S^T| = 1$, since S is orthogonal. The integral becomes

$$I_n = e^c \prod_{i=1}^{n} \int_{-\infty}^{\infty} \exp\{-d_i y_i^2\}\, dy_i = e^c \sqrt{\frac{\pi^n}{\det M}} \tag{B.4}$$

with the determiniant $\det M = \det D = \prod_{i=1}^n d_i$.

For $\vec{\lambda} \in \mathbb{R}^n \setminus \{\vec{0}\}$ we again complete the square to get to a familiar Gaussian expression. Like before, we perform the variable transformation $\vec{x} \mapsto \vec{y} = S\vec{x}$ resulting in the integral

$$I_n = e^c \int_{-\infty}^{\infty} \exp\Big\{-\vec{y}^T D\vec{y} + \vec{\lambda}^T S^T \vec{y}\Big\}\, dy_1 \cdots dy_n$$

Since $S_{ij}^T = S_{ji}$, completing the square in the exponent yields

$$\begin{aligned}
-\vec{y}^T D\vec{y} + \vec{\lambda}^T S^T \vec{y} &= \sum_{i=1}^{n} d_i y_i^2 - \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{n} S_{ij} y_i \\
&= \sum_{i=1}^{n} \left(d_i y_i^2 - y_i \sum_{j=1}^{n} \lambda_j S_{ij} \right) \\
&= \sum_{i=1}^{n} \left[d_i \left(y_i + \frac{1}{2d_i} \sum_{j=1}^{n} \lambda_j S_{ij} \right)^2 - \frac{1}{4d_i} \left(\sum_{j=1}^{n} \lambda_j S_{ij} \right)^2 \right].
\end{aligned}$$

If this is inserted back into the integral, we are left with an n-product of one-dimensional Gaussian integrals

$$I_n = e^c \prod_{i=1}^{n} \int_{-\infty}^{\infty} \exp\left\{ d_i \left(y_i + \frac{1}{2d_i} \sum_{j=1}^{n} \lambda_j S_{ij} \right)^2 - \frac{1}{4d_i} \left(\sum_{j=1}^{n} \lambda_j S_{ij} \right)^2 \right\} dy_i,$$

which leads to

$$I_n = e^c \sqrt{\frac{\pi^n}{\det M}} \underbrace{\exp\left\{ \sum_{i=1}^{n} -\frac{1}{4d_i} \left(\sum_{j=1}^{n} \lambda_j S_{ij} \right)^2 \right\}}_{= \exp\left(\frac{1}{4}\vec{\lambda}^T M^{-1} \vec{\lambda}\right)}. \tag{B.5}$$

Acknowledgements

Thank you to my supervisor Fabian Hassler for taking time every week to have a little fun with stochastics. I am grateful for the close supervision. A special thank you also to Lisa Arndt for lots of Zoom calls, inspiration, and for quick and thorough feedback on my writing.

I would furthermore like to thank my friends and family for their support, in particular Anna, Georg, Gregor and Tobi.

Bibliography

[1] Mikio Namiki. *Stochastic quantization*. Vol. 9. Springer Science & Business Media, 2008.

[2] Georgio Parisi, Yong Shi Wu, et al. "Perturbation theory without gauge fixing". In: *Sci. Sin* 24.4 (1981), pp. 483–496.

[3] Christiane Tretter. *Analysis I*. Mathematik Kompakt. Birkhäuser, 2013.

[4] Brett Williams. "On Stochastic Differential Equations in the Ito and in the Stratonovich Sense". PhD thesis. 2012.

[5] Albert Einstein. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen (German) [On the Movement of Small Particles Suspended in Stationary Liquids Required by the Molecular-Kinetic Theory of Heat]". In: *Annalen der physik* 4 (1905).

[6] Achim Klenke. *Wahrscheinlichkeitstheorie (German) [Probability Theory]*. Vol. 1. Springer, 2006.

[7] John Hull et al. *Options, Futures and Other Derivatives*. Upper Saddle River, NJ: Prentice Hall, 2009.

[8] Fabian Hassler. *Quantenmechanik (German) [Quantum Mechanics]*. 2019.

[9] Mikio Nakahara. *Geometry, topology and physics*. CRC Press, 2003.

[10] Udo Kamps and Erhard Cramer. *Grundzüge der Stochastik (German) [Basics of stochastics]*. Institut für Statistik und Wirtschaftsmathematik, 2018.

[11] Heinz Bauer. *Wahrscheinlichkeitstheorie*. Walter de Gruyter, 1990.

[12] Lawrence C Evans. *An introduction to stochastic differential equations*. Vol. 82. American Mathematical Soc., 2012.