Pro gradu –tutkielma

MATHEMATICAL ASPECTS OF FUNCTIONAL INTEGRATION

Aku Valtakoski

22.11.2000

Ohjaaja: Christofer Cronström

Tarkastajat: Christofer Cronström
Keijo Kajantie

HELSINGIN YLIOPISTO
FYSIIKAN LAITOS
TEOREETTISEN FYSIIKAN OSASTO

PL 9 (Siltavuorenpenger 20 D)
00014 Helsingin Yliopisto

Tiedekunta/Osasto — Fakultet/Sektion — Faculty
Faculty of Science

Laitos — Institution — Department
Department of Physics

Tekijä — Författare — Author
Aku Valtakoski

Työn nimi — Arbetets titel — Title
Mathematical Aspects of Functional Integration

Oppiaine — Läroämne — Subject
Theoretical physics

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Master's thesis	November 2000	64

Tiivistelmä — Referat — Abstract

In the last fifty years functional integration has become an important tool in physics and mathematics. Functional integration techniques are very versatile and can be applied to various types of physical problems. Functional integration is also more than just a calculational method; for example, the Feynman path integral is in fact an independent formulation of quantum mechanics. Functional integration divides into two subclasses: the Wiener integral encountered in connection with diffusion and the Feynman path integral which is used to describe quantum phenomena. In this work we review the properties of the two integrals and compare them to each other. We find that even though they have almost the same mathematical structure the differences between them are profound.

Special emphasis is given to the Feynman path integral. We make the observation that it is not a true integral over a space of functions like the Wiener integral. It is only a shorthand notation for a limit of multiple integrals.

There have been several attempts to formulate a definition of Feynman path integrals that would be mathematically sound. Some of these are reviewed in this work. Even though they have had some success, none of them has achieved the intuitiveness of the original definition by Feynman. We also find them often too abstract to be useful. More research should thus be aimed at finding a proper mathematical definition for Feynman path integrals. This is prompted by their widespread use which often neglects the problems of the definition.

In addition to the discussion on the justification of functional integration we discuss some more specialized subjects, such as stochastic integration, discretization procedure of the Feynman path integral and the Feynman path integral on spaces with curvature.

We also present an extensive list of references on the subject.

Avainsanat — Nyckelord — Keywords
Functional integration, stochastic integral, Wiener integral, Feynman path integral

Säilytyspaikka — Förvaringsställe — Where deposited
Library of Physics, University of Helsinki

Muita tietoja — övriga uppgifter — Additional information

Tiedekunta/Osasto — Fakultet/Sektion — Faculty
Matemaattis-luonnontieteellinen

Laitos — Institution — Department
Fysiikan laitos

Tekijä — Författare — Author
Aku Valtakoski

Työn nimi — Arbetets titel — Title
Mathematical Aspects of Functional Integration

Oppiaine — Läroämne — Subject
Teoreettinen fysiikka

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Pro gradu	Marraskuu 2000	64

Tiivistelmä — Referat — Abstract

Viimeisen viidenkymmenen vuoden aikana funktionaali-integroinnista on tullut tärkeä fysiikan ja matematiikan työkalu. Funktionaali-integrointia voi soveltaa hyvin erilaisiin fysikaalisiin ongelmiin. Tämän lisäksi sillä on muutakin merkitystä; yksi mahdollisuus formuloida kvanttimekaniikka on Feynmanin polkuintegraali.

Funktionaali-integraalit jakaantuvat kahteen aliluokkaan: diffuusion yhteydessä tavattavaan Wienerin integraaliin ja kvanttimekaniikan Feynmanin polkuintegraaliin. Tässä työssä käydään läpi näiden kahden integraalin ominaisuudet ja vertaillaan näitä keskenään. Selviää, että vaikka integraaleilla on lähes samanlainen matemaattinen rakenne on niiden välillä huomattavia eroja.

Erityisesti tarkastellaan Feynmanin polkuintegraalia. Huomataan, että kyseessä ei ole todellinen integraali yli funktioavaruuden vaan ainoastaan merkintä äärellisulotteisen integraalin raja-arvolle. Koska Feynmanin polkuintegraalien määritelmä ei ole matemaattisesti hyvin perusteltu, ovat monet yrittäneet muuttaa määritelmää siten, että se olisi matemaattisesti täsmällinen. Tässä työssä esitellään muutamia tällaisia vaihtoehtoisia määritelmiä. Vaikka ne ovatkin matemaattisesti paremmin perusteltuja, ovat nämä määritelmät usein liian abstrakteja ollakseen yhtä käyttökelpoisia kuin Feynmanin alkuperäinen määritelmä.

Koska polkuintegraaleja käytetään tänä päivänä yleisesti lähes jokaisella fysiikan alalla – usein yllämainitut matemaattiset ongelmat unohtaen – tulisi Feynmanin polkuintegraalien matemaattista määritelmää edelleen tutkia.

Määritelmien lisäksi työssä käsitellään myös erikoistuneempia aiheita, kuten stokastista integrointia, Feynmanin polkuintegraalien diskretisointia sekä polkuintegraaleja kaarevilla avaruuksilla.

Työhön sisältyy lisäksi laaja lista viitteitä aihetta käsitteleviin artikkeleihin ja kirjoihin.

Avainsanat — Nyckelord — Keywords
Funktionaali-integrointi, stokastinen integraali, Wienerin integraali, Feynmanin polkuintegraali

Säilytyspaikka — Förvaringsställe — Where deposited
Fysiikan laitoksen kirjasto

Muita tietoja — övriga uppgifter — Additional information

Contents

Chapter 1

Introduction

Functional integration, also known as path integration or Wiener integration, has become a common tool in physics as well as in some branches of mathematics such as functional analysis and partial differential equations during the last fifty years. It connects the theory of measures and integration with stochastics, particularily with stochastic differental equations.

Functional integration is the theory of integration on spaces of functions. It can also be understood as the extension of ordinary integration theory to spaces of infinite dimensions. Since Lebesque integration theory does not work in such spaces functional integration is highly nontrivial.

Since Feynman [Fey48] introduced path integration methods to physics over fifty years ago functional integration has proved out to be a valuable tool – both in theoretical considerations and in numerical calculations. Quantum field theory, statistical physics as well as analysis of stochastic processes have greatly benefited from the development of functional integration. By "path integration" we shall mean this functional integral of quantum mechanics; the Wiener integral means the functional integral of Brownian motion.

The widespread use of path integration has been the motivation for mathematical research on functional integration in a more general terms. This research has revealed profound connections between path integrals in quantum physics and the Wiener integral applied to Brownian motion. An early review of of the definitions and properties of Feynman's path integrals was made by Gel'fand and Yaglom [G&Y60].

Today path integration has developed to the level that we have a path integral solution to every quantum mechanical problem solvable by the Schrödinger equation. A good review of the current available calculational techniques of path integrals is given by Grosche and Steiner [G&S95]. The paper also contains a large amount of references which touch almost every relevant aspect of functional integration.

Unfortunately most physicist take a very pragmatic view on functional integration and use it as a fool-proof tool. They tend to forget – or neglect – the mathematical subtleties actually involved in the definition and methods of functional integration. Besides that, the analysis of functional integration offers more information than just calculational tools; as mentioned in the first paragraph, the functional integral is also an important mathematical abstraction.

In fact the mathematical foundation of Feynman's path integrals has never been soundly established. The alternating nature of e^{iS} in the Feynman "measure" effectively prohibits the use of well-founded methods of measure theory. In fact, the Feynman "measure" isn't really a measure at all – at least not in the sense of probability theory.

The aim of this work is to give a thorough, but not entirely concise review of the mathematical background of functional integration and, in particular, to show why Feynman path integrals are not mathematically justified. To achieve this, we will review the mathematical properties of Brownian motion and compare them to the properties of Feynman path integrals. We will also study what has been done to correct the lack of mathematical soundness in functional integration.

Although the context of this work is mathematical we will avoid most of complex details. This choice is made at the expense of mathematical completeness but the number of actual mathematical concepts and details involved is so large that it would considerably add to the length of this work. We also wish to keep the material accessible to physicists without an extensive background in mathematics.

We do not not attempt to give a concise review of functional integration or its history. Such information can be found in many textbooks, such as Feynman's own book on the subject with Hibbs [F&H65] or more recent material such as Kleinert's book [Kle90]. Brief definitions of Brownian motion and Feynman path integrals are included as well as some mathematical methods wherever they have some connection to the underlying problems or ambiguities. We shall not go through the calculational methods of functional integration. We shall also not discuss whether or not we can actually explicitely calculate the value of the integral in a closed form.

In chapter 2 we shall study Brownian motion and its mathematical properties, stochastic integration and the Wiener measure and integral. We shall point out the ambiguities that lie in the well-studied theory of Brownian motion and briefly study the implications due to this arbitrariness in the discretization of the Wiener integral. The analysis of Brownian motion also serves as a stepping stone as we proceed to take a closer look at the Feynman path integrals.

Chapter 3 considers the mathematical aspects of Feynman path integrals. We shall study the basic properties of path integrals and find the problems that make them mathematically vague. A review of some important alternative definitions

for Feynman path integrals will be made and these definitions will be put under scrutiny. The discretization procedure used in calculation of path integrals and the solution of path integrals on spaces with curvature, which both have their own mathematical problems are also discussed in this chapter.

1.1 On notation and conventions

- Throughout this work we shall put $c = m = \hbar \equiv 1$ except where explicitely noted. This can be achieved by appropriate choice of the units for time, mass and length. These units are mainly chosen for typographical reasons. Since we are mostly dealing with mathematical rather than physical problems no significant harm is done by doing so.
- Quantum mechanical operator corresponding to a classical quantity A will be denoted as $\widehat{A}$.
- For simplicity we will only consider one-dimensional configuration space and two-dimensional phase space. Generalizations to higher dimensions and other types of spaces are (usually) straightforward. Wherever necessary, these generalizations will be analyzed seperately. This will be especially done in section 3.5, where we consider functional integration on non-eucledian spaces.
- Variable x will generally be the coordinate in configuration space, whereas variables q and p are reserved for coordinate and momentum, respectively, in phase space.
- x_i is shorthand notation for $x(t_i)$. Note that this applies also to other quantities besides the coordinate.
- $\mathbb{R}^+ \doteq [0, \infty]$

1.2 Acknowledgements

I would like to thank professor Christofer Cronström for his patience and his interest in even the tiniest details and Heikki and Hilma Honkanen Foundation for supporting my studies financially.

Chapter 2

Brownian Motion and the Wiener Integral

Brownian motion is named after the Scottish botanist Robert Brown, who first observed the irregular motion of pollen grains in a liquid. Later this motion was studied in detail by Einstein (a collection of his papers concerning Brownian motion can be found in [Ein56]) and Wiener ([Wie23],[Wie24],[Wie30]), after whom we now call the functional integral that arises from the mathematical model of this random motion.

The physical model of Brownian motion is as follows: a particle (in every sense macroscopical and therefore classical) moves in a medium which projects a random force on the particle. This force causes the particle to move on very irregular and nondeterministic paths. Furthermore, these paths are now known to have a fractal nature. As we shall see, this fractal behaviour of Brownian motion is the most important feature that affects its mathematical analysis as well as that of Feynman path integrals.

Since the paths of a Brownian particle are nondeterministic, we must analyze it with the machinery of probability theory and stochastic processes. From a physicist's point of view, this corresponds to statistical physics. Simply expressed, we can only know the probabilities for the particle to move from point a to point b. The exact path of the particle cannot *a priori* be known .

In what follows we shall take a mathematical point of view and disregard most of the physics involved in describing Brownian motion. The definition of Brownian motion used here is not the usual one found in books on functional integration. This is done because we want to stress the fact that one arrives at the same conclusions about Brownian motion even though there are different ways to define it.

We shall closely follow Øksendal [Øks95] in defining the important concepts of probability space and stochastic processes and also later when we explore the

properties of Brownian motion. This choice is motivated by Øksendal's similar approach to the subject. Unfortunately we must omit even some core material in order to keep this work compact and not to lead the discussion astray. For more precise and deeper information one should consult Øksendal's book and references therein.

2.1 Definitions

Definition 2.1 (Probability space) *The triple* $(\Omega, \mathcal{A}, P)$ *is called a* probability space, *if* $\mathcal{A}$ *is a* σ*-algebra on the set* Ω *and* P *is measurable function* $P : \mathcal{A} \to [0,1]$ *on the measurable space* $(\Omega, \mathcal{A})$ *so that*

1. $P(\emptyset) = 0$, $P(\Omega) = 1$ *and*
2. $A_1, A_2, \ldots \in \mathcal{A}$ *and they are disjoint (i.e.* $A_i \cap A_j = \emptyset$ *if* $i \neq j$*)*

As one can see, the notion of probability is very closely related to the theory of measures and integration. Therefore we are equipped with the powerful machinery of integration theory when we tackle problems in probability and Brownian motion. Many important results concerning Brownian motion have been derived by Kolmogorov (see, for example, his own book on the subject [Kol56]).

A random variable X is simply a measurable function $X : \Omega \to \mathbb{R}$. ($\mathbb{R}$ could, of course, also be some other space.) An important notion in probability theory is *almost sureness*. We denote this by

$$X \overset{a.s.}{=} Y$$

if

$$P\{\omega | X(\omega) \neq Y(\omega)\} = 0.$$

Or, in expressed in words, X and Y differ from each other only on a set of measure zero. Note that this set is not necessarily an empty set.

Definition 2.2 (Independence) *Two subsets ("events")* $A, B \in \mathcal{A}$ *are independent if*

$$P(A \cap B) = P(A)P(B).$$

Definition 2.3 (Stochastic process) *A stochastic process is a parametrized collection of random variables*

$$\{X_t\}_{t \in T}$$

defined on a probability space $(\Omega, \mathcal{A}, P)$ *and assuming values in* $\mathbb{R}$.

(The restriction to one dimension is made only because of our decision to simplify notation. The space of values taken by the stochastic process is usually more general – for example $\mathbb{R}^n$.)

Notice that a stochastic process is a function $X(\omega,t) : \Omega \times \mathbb{R}^+ \to \mathbb{R}$ (t is usually taken positive). Therefore, if we fix $\omega \in \Omega$ we get the function

$$t \to X_t(\omega); t \in \mathbb{R}^+$$

which is called the *path* of X_t. This can be thought of as the "path" of one Brownian particle in the liquid or as one experiment (if we think in the context of quantum mechanics).

Also, if we fix t we see that X is a mapping $X : \Omega \to \mathbb{R}$ for all fixed t. Thus, X can take any possible value in $\mathbb{R}$ depending on the process. From our viewpoint we can then consider Ω as a space of paths, parametrized by elements $\omega \in \Omega$.

If $\{X_t\}$ and $\{Y_t\}$ are stochastic processes so that

$$P(\{\omega; X_t(\omega) = Y_t(\omega)\}) = 1 \quad \forall t$$

then $\{X_t\}$ is called a *version* of $\{Y_t\}$. Simply put the two processes are indistiguishable in terms of probability distributions. However, it should be noted that their path properties and can be different. For example, there can be a denumerable set of points where X_t and Y_t have different value.

Definition 2.4 (Brownian motion) *Let X_t be a stochastic process starting at $X_0 = 0$ with the following properties:*

1. *The process X_t has normal distribution with mean 0 and variance t*
2. *The future values of X_t are independent of all X_s, $s < t$*
3. *The increments $X_i - X_j$, $X_j - X_k$ are independent of each other*

Then the stochastic process considered is called standard Brownian motion *(hereafter simply Brownian motion).*

From this definition one can readily explicitely write down the probability for the particle to move from $(0,0)$ to (x,t):

$$P(0,x,t) = \frac{1}{\sqrt{2\pi t}} \exp\Big(-\frac{x^2}{2t}\Big). \tag{2.1}$$

If the process instead starts at (x_0,t_0), the probability for *transition* is

$$P(x_0,x,t_0,t) = \frac{1}{\sqrt{2\pi(t-t_0)}} \exp\Big(-\frac{(x-x_0)^2}{2(t-t_0)}\Big). \tag{2.2}$$

(Note that we assume that $t \geq t_0$.)

If we integrate over the endpoint in equation (2.1) we arrive at the expected result

$$\int_{-\infty}^{\infty} dx P(x_0, x, t) = 1$$

reassuring us that probability is conserved. One easily sees that both (2.1) and (2.2) are Gaussian probability measures. This holds also for Brownian motion in higher dimensions.

The definition of Brownian motion we used here is not unique; one might start with the above transition probabilities and find that they describe a stochastic process with the properties stated in definition 2.4. Yet another way to construct Brownian motion is to start from the definition of Wiener integrals (see section 2.4). It should be noted that in the end all of these approaches are mathematically equal and lead to same properties for Brownian motion.

It is worthwhile to notice that since Brownian motion does not depend on its history (property number 2), the transition probability will depend only on the difference $(t - t_0)$. In fact, when we look at equation (2.2) we see that it depends on the coordinates only through the differences $(x - x_0)$. Brownian motion is thus homogenious in both time and space and is an example of a *Markovian process*. Further properties of Brownian motion, especially of its paths, will be discussed in the next section.

2.2 Properties of Brownian motion

Using the probability measure (or distribution) defined in equation (2.1) we define *expectation* in the normal way as

$$E[f] := \int_{\mathbb{R}} f(x) d\mu(x) = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} dx\, f(x) e^{\frac{(x-x_0)^2}{2t}}$$

where f is a Borel measurable integrable function and $d\mu(x)$ is the probability measure, in this case the measure of equation (2.1).

Since we are going to discuss the properties of Brownian motion, it is very useful to calculate the characteristic function of $(x - x_0)$, which defined as

$$\phi(u) := E[e^{iu(x-x_0)}]. \tag{2.3}$$

If we write it down explicitly we find that

$$E[e^{iu(x-x_0)}] = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} dx e^{iu(x-x_0)} e^{\frac{(x-x_0)^2}{2t}} = e^{iux_0} e^{-\frac{1}{2}u^2 t}$$

Having calculated the characteristic function, all the moments of $(x-x_0)$ can be found by differentiation:

$$E[(x-x_0)^n)] = (-i)^n \frac{d^n \phi}{du^n}|_{u=0} \tag{2.4}$$

We find that for all even n the moment is zero. For $n=2$ we find the variance of Brownian motion. It is

$$E[(x-x_0)^2] = t, \tag{2.5}$$

just as we would expect on the basis of the definition 2.4. Furthermore, if we calculate the quartic moment ($n=4$) we have the result

$$E[(x-x_0)^4)] = 3t^2. \tag{2.6}$$

This result is important since we can now apply Kolmogorov's continuity theorem (see, for example, Øksendal [Øks95]), which states that the paths of Brownian motion are continous. (Or, more exactly, there exists a continous version of x.) A physical argument can also be used to "prove" the continuity of the paths; discontinuous paths are clearly unphysical – the Brownian particle does not make "jumps".

One might wonder what happens if we take $t \to 0$ in equation (2.1). It seems as if the probability would blow up as we take the limit because of the term $\exp(-(x-x_0)^2/2t)$. But, if we take Fourier transformation of $P(x_0,x,t)$ (actually we are calculating the characteristic function again) we get the result

$$\mathcal{F}\{P(x_0,x,t)\} := e^{iux_0} e^{-\frac{1}{2}u^2 t}.$$

If we then take $t \to 0$, we see that the Fourier transform is equal to the Fourier transform of Dirac's δ-function. Thus we conclude that

$$\lim_{t\to 0} P(x_0,x,t) = \delta(x-x_0).$$

As we have assured ourselves that the Brownian motion has continous paths, it seems appropriate that we say something about the differentiability of these paths. It can be proved (see, for example Breiman [Bre68]) that **the paths are nondifferentiable** for almost all ω. This is the most important property of Brownian motion, and can be expressed by saying that the paths have a *fractal dimension* larger than 1.

In the end of this section we recall that Browian motion is just a single example of a stochastic process. There are numerous other processes satisfying the same basic definition of a stochastic process with otherwise different properties. In this work we will mostly consider continous stochastic processes. We mention that in addition to them there are also many important stochastic processes with discrete time dependence.

2.3 Stochastic integration

Let us now return to strictly mathematical constructs involving stochastic processes. In this section we familiarize ourselves with the concept of stochastic integration and see how the "$(dx)^2 = dt$" property affects this procedure. We will also see that there are two (actually, infinite number of) different ways to actually calculate a stochastic integral yielding a different result.

This brief excursion to stochastics is motivated by the fact that we will learn how Brownian motion behaves under integration and more about the path properties of Brownian motion.

Stochastic integration is required whenever one encounters *stochastic differential equations* (or SDEs). These are usually normal differential equations where a random element is added. (For example, to model the size of population for some species.) A major subclass of these differential equations are of the form

$$\frac{dX}{dt} = b(t, X_t) + \sigma(t, X_t) \cdot \text{“noise”}, \tag{2.7}$$

where b and σ are some given functions. The "noise" term needs a proper mathematical interpretation. Let us represent it as a stochastic process W_t. We assume that W_t has the following properties:

(i) $t_1 \neq t_2 \Rightarrow W_{t_1}$ and W_{t_2} are independent

(ii) $\{W_t\}$ is stationary i.e. the distribution of W_{k+t} is independent of t for all $k > 0$.

(iii) $E[W_t] = 0$ for all t.

Unfortunately, there is no "reasonable" stochastic process satisfying properties (i) and (ii). Such a W_t cannot have continous paths. It is, however, possible to represent W_t as generalized stochastic process called the *white noise process*. We shall construct such process by first considering the discrete version of equation (2.7):

$$X_{k+1} - X_k = b(t_k, X_k)\Delta t_k + \sigma(t_k, X_k)W_k\Delta t_k, \tag{2.8}$$

where $\Delta t_k := t_{k+1} - t_k$. The next thing to do is to replace $W_k\Delta t_k$ by $\Delta B_k = B_{k+1} - B_k$, where $\{B_t\}$ is some suitable stochastic process. It turns out that $\{B_k\}$ is in fact a Brownian motion (that's why we chose the letter B). With this notation, we can obtain from equation (2.8) by summation:

$$X_k = X_0 + \sum_{j=0}^{k-1} b(t_j, X_j)\Delta t_j + \sum_{j=0}^{k-1} \sigma(t_j, X_j)\Delta B_j. \tag{2.9}$$

Equation (2.9) immediately raises the question: What happens if we take $\Delta t \to 0$? If the limit exists, we would be able to write the limit in terms of usual integration notation as

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s \tag{2.10}$$

Especially the last term,

$$\int_0^t \sigma(s, X_s)dB_s \tag{2.11}$$

needs some clarifying and special treatment, for we are defining integration with respect to Brownian motion! If we are able to define the integral it means that $X_t = X_t(\omega)$ of equation (2.10) is a stochastic process and a solution to the original stochastic differential equation (2.7).

In order to define integrals of the form (2.11) we apply the usual method of probability theory: first we define it on simple functions i.e. functions that are piecewise constant. Simple functions are of the form

$$\phi(t, \omega) = \sum_{j \geq 0} e_j(\omega) \cdot \chi_{[j2^{-n}, (j+1)2^{-n})}(t), \tag{2.12}$$

where χ denotes the characteristic (or indicator) function and n is a (large) natural number.

We could then generalize this result to apply to non-negative functions and finally to arbitrary measurable functions. However, this procedure is not essential to our work and therefore we shall present only the main points of this definition procedure.

For functions like (2.12) we can define

$$\int_S^T \phi(t, \omega)dB_t(\omega) = \sum_{j \geq 0} e_j(\omega)[B_{t_{j+1}} - B_{t_j}](\omega). \tag{2.13}$$

Let us now choose

$$\begin{aligned} \phi_1(t, \omega) &= \sum_{j \geq 0} B_{j2^{-n}}(\omega) \cdot \chi_{[j2^{-n}, (j+1)2^{-n})}(t) \\ \phi_2(t, \omega) &= \sum_{j \geq 0} B_{(j+1)2^{-n}}(\omega) \cdot \chi_{[j2^{-n}, (j+1)2^{-n})}(t) \end{aligned}$$

It seems that these two functions are very similar approximations to

$$f(t, \omega) = B_t(\omega).$$

But, if we calculate their expectations of their integrals over some time interval $[0,T]$, we get

$$E[\int_0^T \phi_1(t,\omega)dB_t(\omega)] = \sum_{j\geq 0} E[B_{t_j}(B_{t_{j+1}} - B_{t_j})] = 0,$$

since $\{B_t\}$ has independent increments. For ϕ_2 the result is different,

$$\begin{aligned} E[\int_0^T \phi_2(t,\omega)dB_t(\omega)] &= \sum_{j\geq 0} E[B_{t_{j+1}}(B_{t_{j+1}} - B_{t_j})] \\ &= \sum_{j\geq 0} E[(B_{t_{j+1}} - B_{t_j})^2] = T, \end{aligned}$$

by equation (2.5).

The two superficially equal approximations thus give different values for the integral. Note also that this result does not depend on how large n we have chosen.

This result is the first indication we encounter to display the basic fact that the paths of Brownian motion are too jagged – their total variation on interval $[0,T]$,

$$V(B(t)) = \sup_n \Big[\sum_{j=1}^n |B(t_j) - B(t_{j-1})| \Big], \quad 0 = t_0 < t_1 < \ldots < t_n = T,$$

is infinite. This prohibits us from defining the integral (2.10) in the usual Riemann-Stieltjes sense.

However, we proceed to define the general integral 2.10 as follows:

Definition 2.5 (Stochastic integral) *Let $f(t,\omega)$ be a suitable function (square integrable with respect to t and measurable with respect to Brownian motion B_t). The stochastic integral*

$$\int_S^T f(t,\omega)dB_t(\omega) \tag{2.14}$$

is then equal to the limit

$$\lim_{n\to\infty} \sum_{j\geq 0} f(t_j^*,\omega)[B_{t_{j+1}} - B_{t_j}],$$

where the points t_j^* belong to the intervals $[t_j, t_{j+1}]$.

Notice that in order to find a value of a stochastic integral one must define the points t_j^* – i.e. define the discretization $f(t,\omega)$. As we have seen, this choice directly affects the value of the integral. The two following choices for t_j^* have turned out to be the most useful ones:

1. $t_j^* = t_j$, the left end point of the interval $[t_j, t_{j+1}]$. This is referred to as the *prepoint prescription*. This choice leads to the *Ito integral*, which we hereafter denote by
$$\int_S^T f(t,\omega) \bullet dB_t(\omega)$$
2. $t_j^* = (t_j + t_{j+1})/2$, the mid point of the interval $[t_j, t_{j+1}]$. This is also known as the *midpoint prescription*. This choice leads to the *Stratonovich integral*, denoted by
$$\int_S^T f(t,\omega) \circ dB_t(\omega)$$

Some important properties of both Ito and Stratonovich integrals will be discussed in the two following subsections.

2.3.1 Ito integrals

As noted above, Ito integrals are produced by the choice $t_j^* = t_j$ as the discretization rule. The choice gives Ito integrals the property of "not looking into the future" i.e. on the interval $[t_j, t_{j+1}]$ the only value that matters is B_t. This property is most profoundly expressed in terms of *martingales*. But first we define what we mean by conditional expectation.

Definition 2.6 (Conditional expectation) *Let $\mathcal{H} \subset \mathcal{A}$ and X be a random variable with $E|X| < \infty$. The conditional expectation $E[X|\mathcal{H}]$ is the function $\Omega \to \mathbb{R}$ with the following properties:*

(i) $E[X|\mathcal{H}]$ *is $\mathcal{H}$-measurable*

(ii) $\int_H E[X|\mathcal{H}]dP = \int_H X\, dP$ *for all* $H \in \mathcal{H}$

Note that $E[X|\mathcal{H}]$ is a function and not a number like the normal expectation. Conditional expectation is a generalization of the conditional probability in simple probability theory:

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

where $A, B \subset \Omega$.

Now we can define

Definition 2.7 (Martingale) *Let $\{M_t\}_{t\geq 0}$ be a stochastic process on the probability space $(\Omega, \mathcal{A}, P)$ such that*

(i) M_t *is* $\mathcal{F}_t$*-measurable for all* t

(ii) $E[|M_t|] < \infty$ *for all* t

(iii) $E[M_s|\mathcal{F}_t] = M_t$ *for all* $s \geq t$

where $\{\mathcal{F}\}_{t\geq 0}$ is an increasing family of σ-algebras on Ω such that

$$0 \leq s < t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{A}.$$

(Such a family is called a filtration*.) The stochastic process is then called a* martingale*.*

Let us now check that Ito integrals really have the properties required of a martingale with respect to the σ-algebras generated by the Brownian motion $\{B_s; s \leq t\}$:

(i) $E[\int_S^T f(t,\omega) \bullet dB_t(\omega)] = 0 < \infty$

(ii) $\int_S^T f(t,\omega) \bullet dB_t(\omega)$ is $\mathcal{F}_T$-measurable.

(iii) $E[B_s|\mathcal{F}_t] = E[B_s - B_t + B_t|\mathcal{F}_t] = E[B_s - B_t|\mathcal{F}_t] + E[B_t|\mathcal{F}_t] = 0 + B_t = B_t$

The property 1 follows from taking first a simple function $f_n(t,\omega)$ and proving the equality and then taking the limit $n \to \infty$. Property 2 can be proved in the same fashion. In 3 we have used the fact that $B_s - B_t$ is independent of $\mathcal{F}_t$ and $E[B_t|\mathcal{F}_t] = B_t$ since B_t is $\mathcal{F}_t$-measurable. For this fact and others concerning conditional probability one may consult Williams' book [Wil91], for example.

Thus we have proved that Ito integrals are martingales. Martingales have many "nice" properties and the theory of martingales is well-developed. We are thus equipped with lots of useful concepts and theorems, including the important Doob's martingale inequality. The proof of the theorem and numerous other results for martingales can be found in Williams' book, which takes martingales as a starting point and, using martingales, develops the structure of probability theory.

An equally important fact – for example, in light of applications – is the *martingale representation theorem*. It states that any martingale with respect to the filtration $\{\mathcal{F}\}$ generated by Brownian motion can be represented as an Ito integral. For proof see Øksendals book [Øks95].

Let us now prove that for Ito integrals

$$\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t, \tag{2.15}$$

assuming that $B_0 = 0$.

First we put $\phi_n(s,\omega) = \sum B_j(\omega) \cdot \chi_{[t_j,t_{j+1})}(s)$. Then

$$E\big[\int_0^t (\phi_n - B_s)^2 ds\big] = E\big[\sum_j \int_{t_j}^{t_{j+1}} (B_j - B_s)^2 ds\big]$$
$$= \sum_j \int_{t_j}^{t_{j+1}} (s - t_j) ds = \sum_j \frac{1}{2}(t_{j+1} - t_j)^2 \to 0 \text{ as } \Delta t_j \to 0$$

So

$$\int_0^t B_s dB_s = \lim_{\Delta t_j \to 0} \int_0^t \phi_n dB_s = \lim_{\Delta t_j \to 0} \sum_j B_j \Delta B_j.$$

Now observe that

$$\Delta(B_j^2) = B_{j+1}^2 - B_j^2 = (B_{j+1} - B_j)^2 + 2B_j(B_{j+1} - B_j)$$
$$= (\Delta B_j)^2 + 2B_j \Delta B_j,$$

and therefore

$$B_t^2 = \sum_j \Delta(B_j^2) = \sum_j (\Delta B_j)^2 + 2\sum_j B_j \Delta B_j$$

Dividing equation (2.3.1) by 2, moving the second term on the right hand side to the left side of the equation and using the result $\sum_j (\Delta B_j)^2 \to t$ in $L^2(P)$ (in the mean square sense) as $\Delta t_j \to 0$ we arrive at equation (2.15).

The above calculation serves as a general example of how Ito integrals are calculated and also as an example of another important feature of Ito integrals: they do not behave as ordinary integrals. For example, the term $-\frac{1}{2}t$ would not appear in normal integration. This, in fact, is a clear demonstration of the property "$(dX)^2 \sim dt$" of Brownian motion.

Furthermore, if we write the equation (2.15) as

$$\frac{1}{2}B_t^2 = \frac{1}{2}t + \int_0^t B_s dB_s$$

and consider it as a mapping $g(x) = \frac{1}{2}x^2$ of the Ito integral $B_t = \int_0^t dB_s$, we see that the result is not of the form $\int_0^t f dB_s$. This shows that normal rules of changing variables do not hold when we calculate Ito integrals. Instead, we now have

the result that *Ito processes* $dX_t = u(t,B_t)dt + v(t,B_t)dB_t$ form a closed group; any mapping of X_t is again an integral of the form (2.10).

The above result and many others explicit forms of Ito integrals can be obtained by a formula corresponding to the chain rule of differentiation, the *Ito formula*. It can be stated shortly as (note that this definition applies only to one-dimensional Brownian motion)

Theorem 2.1 (The Ito Formula) *Let X_t be an stochastic process given by the SDE*

$$dX_t = udt + vdB_t,$$

Where u is almost surely integrable and v is almost surely square integrable with respect to the probability measure P on Ω. Let $g(t,x)$ be a twice continously differentiable on $\mathbb{R}^+ \times \mathbb{R}$. Then

$$Y_t = g(t,X_t)$$

is again a stochastic process, and the differential of Y_t is

$$dY_t = \frac{\partial g}{\partial t}dt + \frac{\partial g}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(dX_t)^2, \tag{2.16}$$

where all partial differentials of g are computed at the point (t,X_t) and the term $(dX_t)^2$ is computed according to the rules

$$dt \cdot dt = dt \cdot dB_t = dB_t \cdot dt = 0, \quad dB_t \cdot dB_t = dt. \tag{2.17}$$

Equation (2.17) explicitely shows that $(dX_t)^2 = dt$. We will not prove the Ito formula here but rather refer to Øksendal's book [Øks95]. One can easily see that by choosing $X_t = B_t$ and $g(t,x) = \frac{1}{2}x^2$ and using the Ito formula one arrives at the result (2.15). Other results of stochastic integrals, for example the integration-by-parts rule for Ito integrals, can also be derived by using the Ito formula.

2.3.2 Stratonovich integrals

If we choose $t_j^* = (t_j + t_{j+1})/2$ instead of t_j we arrive at the Stratonovich interpretation of the integral

$$\int_0^t f(s,\omega)dB_s(\omega)$$

There are many reasons to pick such a choice; for example, the Stratonovich integral has the same rule for changing variables as ordinary integrals. This fact

makes it a natural choice when one considers stochastic differential equations on manifolds. We will see this property by calculating the integral

$$\int_0^T B_t \circ dB_t. \tag{2.18}$$

We begin by noticing that

$$\int_0^T f(t,\omega) \circ dB_t(\omega) = \lim_{\Delta t_j \to 0} \sum_j f(t_j^*, \omega)\Delta B_j, \quad \text{where } t_j^* = \frac{1}{2}(t_j + t_{j+1}).$$

Using this, we write equation (2.18) in discretized form:

$$\begin{aligned}
\int_0^T B_t \circ dB_t &= \sum_j B_j^* \Delta B_j \\
&= \sum_j \frac{1}{2}(B_{j+1} + B_j)(B_{j+1} - B_j) \\
&= \sum_j \frac{1}{2}(B_{j+1}^2 - B_j^2) \\
&= \sum_j \frac{1}{2}\Delta(B_j)^2.
\end{aligned} \tag{2.19}$$

Summing over the index j we see that $\sum_j \Delta(B_j)^2 = B_T^2$, actually regardless of the number N of discretization points. Thus, we have computed the integral (2.18) and write the solution as

$$\int_0^T B_t \circ dB_t = \frac{1}{2}B_T^2. \tag{2.20}$$

As one can see from the above calculation, Ito and Stratonovich interpretations of the stochastic integral are generally different from each other. In some cases they do however coincide. To be more precise, this happens whenever the function to be integrated varies "smoothly" enough with t. It can also be proved (see Stratonovich [Str66]) that we can transform these integrals into each other according to the formula

$$\int_0^t \sigma(s, X_s) \circ dB_s = \frac{1}{2}\int_0^t \frac{\partial \sigma(s, X_s)}{\partial x}\sigma(s, X_s)ds + \int_0^t \sigma(s, X_s) \bullet dB_s. \tag{2.21}$$

Because of the explicit connection between the two interpretations it suffices for almost all mathematical purposes to consider only one of them. One can easily then revert back to the other interpretation by using the formula (2.21).

The way we calculated the integral (2.14) with Ito and Stratonovich interpretations demonstrates the different techniques that can be applied when calculating stochastic integrals. The Ito formula is probably the easiest way of doing these integrals, although it is limited to Ito integrals.

The most important lesson to be learnt from this section is the following: *in order to calculate integrals such as equation (2.14) one must also define what interpretation of the integral one uses; i.e. one must implement a certain rule of discretization*. In different situations one choice just makes more sense than the others. This is a phenomenon that will appear again in the context of functional integration – the Wiener integrals and the Feynman path integrals.

2.4 The Wiener measure and integral

In this section we shall return to the original problem of expressing Brownian motion in terms of functional integration. We are still short of an integral defined on a space of functions rather than configuration space.

Consider now that the Brownian particle moves in succession from (x_0,t_0) to (x_1,t_1) and from there on to (x_2,t_2). The probability for movement from (x_0,t_0) to (x_2,t_2) can then be understood as follows: the particle moves to some point x_1 at fixed instant of time t_1. The point can be anywhere in the configuration space. This probability for this is expressed as

$$P(x_0,t_0,x_2,t_2) = \int_{-\infty}^{\infty} dx_1 P(x_0,t_0,x_1,t_1)P(x_1,t_1,x_2,t_2). \tag{2.22}$$

Equation (2.22) is also called the *Einstein-Smoluchowski-Chapman-Kolmogorov equation*.

If we make further restrictions on the path of the particle and demand that it on its way from (x_0,t_0) to (x_k,t_k) it must go through all the points $(x_1,t_1),(x_2,t_2),$ $\ldots,(x_{k-1},t_{k-1})$ we find that the probability for such movement is equal to the product of the probabilities of movements on the required intervals (implied by independency properties of Brownian motion):

$$P(x_0,t_0,x_k,t_k)_{\{x_i\}} = \prod_{i=1}^{k} P(x_i,t_i,x_{i-1},t_{i-1}), \tag{2.23}$$

where $\{x_i\}$ denotes that the path *must* go through the point set beforehand. Even though we limit the path to go through these prescribed points, Brownian motion can take any value between two points i.e. the path needs not to be straight between the points.

If we wish to find the probability for the particle to travel from (x_0, t_0) to (x_k, t_k) through *any* path we must allow the intermediate points to take any applicable value. This is done by integrating over the intermediate points. Explicitely done this yields

$$P(x_0, t_0, x_k, t_k) = \int_{\mathbb{R}\times\cdots\times\mathbb{R}} \prod_{i=1}^{k} dx_i \, P(x_i, t_i, x_{i-1}, t_{i-1}), \tag{2.24}$$

where dx_i is the ordinary Lesbesque measure of $\mathbb{R}$. Notice that the starting point and the end point of the interval are not integrated over, since they are fixed. The above equation can also be understood as the being generated by using the ESKC equation (2.22) recursively.

If we take the set $\{t_1, t_2, \ldots, t_k\}$ to be a finite set (with N elements), we get

$$P(x_0, t_0, x_k, t_k) = \int_{\mathbb{R}\times\cdots\times\mathbb{R}} \prod^{N-1} dx_i \prod^{N} \left(\frac{1}{\sqrt{2\pi(t_i - t_{i-1})}}\right) \exp\left(-\sum^{N} \frac{1}{2} \frac{(x_i - x_{i-1})^2}{(t_i - t_{i-1})}\right). \tag{2.25}$$

(All products and summations start from $i = 1$.)

Now we choose to divide the interval $[t_0, t_k]$ into N equal subintervals, all with length $\varepsilon = (t_k - t_0)/N$. In principle a nonlinear division can be done, but when we take $N \to \infty$ the length of each subinterval will approach zero no matter how the division is done. Thus, $t_i = t_0 + i\varepsilon$, with $t_N = t_k$ and $t_i - t_{i-1} = \varepsilon$

Inserting this division to subintervals into the equation (2.25) we get

$$P(x_0, t_0, x_k, t_k) = \int_{\mathbb{R}\times\cdots\times\mathbb{R}} \prod dx_i \left(\sqrt{\frac{\varepsilon}{2\pi}}\right)^{N-1} \exp\left(-\sum^{N-1} \varepsilon \frac{(x_i - x_{i-1})^2}{2\varepsilon^2}\right).$$

Let us now ponder on what we have just calculated; first we calculated the probability for a certain path in the configuration space. Then we extended this to all possible values at each t_i, keeping at the same time the number of such intervals finite. And as the last step, we let $N \to \infty$. The resulting integral could then be interpreted as a single integral over a space of functions $f : [t_0, t_k] \to \mathbb{R}$ rather than multiple integrals over the underlying configuration space.

We could then write the integral formally as

$$P(x_0, t_0, x_k, t_k) = \int \mathcal{D}x \, \exp\left(-\int_{t_0}^{t_k} dt \, \frac{1}{2}\dot{x}^2\right), \tag{2.26}$$

where $\mathcal{D}x = \lim_{N\to\infty} \left(\sqrt{\varepsilon/2\pi}\right)^N \prod^{N-1} dx_i$ and $\dot{x} = \frac{dx}{dt}$. Note that neither the measure $\prod dx_i$ nor the normalizing constant $\left(\sqrt{\varepsilon/2\pi}\right)^N$ has a limit on its own. However, (2.25) does have a limit as stated above.

The above result is still merely a limit of a discrete product of integrals rather than a integral of its own. To truly establish a functional integral of Brownian motion we need the Kolmogorov extension theorem, which tells us that there is a unique measure – and therefore an integral – on the space of paths in $\mathbb{R}$. This measure coincides with the usual product measure stated above for all values of the number of discretization points. (The measures defined on these points are usually called *cylinder measures*.)

The discovered measure is the *conditional Wiener measure*. The word conditional is used because we have fixed the endpoint of the path. Nevertheless, equation (2.26) is finally mathematically on solid ground and we have the *Wiener integral* in our hands. We will denote the Wiener measure as $\mathcal{D}_W x$.

The implications of Wiener measure are quite interesting. Consider the fact from section 2.2 that the paths of Brownian particles are nondifferentiable. It can be shown that in fact only these nondifferentiable paths have nonzero Wiener measure and thus continous paths do not affect the value of Wiener integrals. Rivers [Riv87] describes this with an interesting analogy. He points out that the situation here is similar to the one with Lebesque measure and rational numbers; although rational numbers are dense in the set of real numbers (as are the continous paths in the space of paths) their Lebesque measure is zero.

We can also see that the Wiener measure of a single path is zero whether it is differentiable or not. This can be seen, for example, by taking equation (2.23) and considering the limit $N \to \infty$. This fact is also in analogy with the Lebesque measure – the measure of a single point in $\mathbb{R}$, or in any denumerable set, is 0.

Note that the derivative $\dot{x}$ is present in equation (2.26). How is this possible if the paths having nonzero measure are almost certainly nondifferentiable? This implies that if a path contributes to the value of the integral, the exponential term must be infinite.

2.4.1 Diffusion

The Wiener integral can also be understood as Green's function of the diffusion equation

$$\frac{\partial \psi}{\partial t} = \frac{1}{2}\frac{\partial^2 \psi}{\partial x^2} \tag{2.27}$$

(subject to initial condition $\psi(x,0) = \delta(x)$). In fact the original Browian motion can also be shown to satisfy the diffusion equation. (Which, in the language of mathematics, is called *Kolmogorov's backward equation*.) Roughly speaking, both of these correspond to the situation in which we have

$$u(t,x) = E[f(B_t)],$$

where E is the expectation with respect to the measure in question and B_t is a Brownian motion. If we are talking about the Wiener measure we replace B_t by a general path $x(t)$. Then u satisfies the equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} \tag{2.28}$$

with $u(0,x) = f(x)$ as the initial condition.

One can easily recognize that the probability measure of Brownian motion (or the Wiener measure) is the Green's function since

$$u(t,x) = \int \mathcal{D}x \; f(x),$$

where $\mathcal{D}$ is the appropriate measure. Note that if $f(x) = \delta(x)$ we have $u(t,x) = \mathcal{D}x$.

One way to construct the Wiener measure and to analyze the properties of Wiener integrals is to view it as the Green's function of the diffusion equation and to assume that the Wiener measure exists. Then one uses Kolmogorov's extension theorem to show that the measure has a unique extension to cylinder measures on $\mathbb{R}$. Doing this, one could derive the properties of transition probability for Brownian motion and the ESKC-equation (2.22). This way of defining Wiener integrals is for example taken by Glimm and Jaffe [G&J81], and is quite opposite to the way we have found Wiener integrals.

Let us now generalize the system; we assume that in addition to diffusion there is a heat sink in the system. This corresponds to adding a potential term $V \cdot u$ to the equation (2.27), where V is a continous function in $\mathbb{R}$ (with possible limitations on the values it may take). The well-known solution of the new equation, also known as the Feynman-Kac formula, is

$$u(t,x) = E[\exp(-\int_0^t V(x(t))ds) f(x(t))], \tag{2.29}$$

expressed here in terms of the Wiener measure. If we write down the expectation explicitely, we arrive at the equation

$$u(t,x) = \int \mathcal{D}_w x \; \exp(-\int_0^t ds[\frac{1}{2}\dot{x}^2 + V(x)]) f(x(t)). \tag{2.30}$$

The term in exponential resembles a familiar construct; it is almost like the classical action integral $\int_0^t dt\mathcal{L}$, where $\mathcal{L} := \frac{1}{2}\dot{x}^2 - V(x)$ is the Lagrangian.

Beside the diffusion equation, Brownian motion and the Wiener integrals can be applied to other partial differential equations as well. They also have other important applications, such as solutions of boundary value problems, for example.

2.4.2 Discretization of the Wiener integral

One of the virtues of equation (2.30) is that it allows easy linear coordinate transformations; it behaves exactly as an ordinary integral under these transformations. We must bear in mind that the paths that contribute to the value of the integral are nondifferentiable. This causes trouble when one tries to apply a nonlinear coordinate transformation, just as in stochastic integration. A fine example of this is a free particle (thus $V = 0$) moving in two dimensions. The transition probability is then

$$P(x,y,t,x_0,y_0,t_0) = \int \mathcal{D}_w x \mathcal{D}_w y \exp(-\int_{t_0}^{t} dt[\frac{1}{2}(\dot{x}^2+\dot{y}^2)]). \tag{2.31}$$

If we would consider the above equation literally, after transforming from Cartesian coordinates to cylinderical coordinates ($x = r\cos\phi,\ y = r\sin\phi$) we would have

$$P'(r,\phi,t,r_0,\phi_0,t_0) = \int \mathcal{D}_w r \mathcal{D}_w \phi\ J[r] \exp(-\int_{t_0}^{t} dt[\frac{1}{2}(\dot{r}^2+r^2\dot{\phi}^2)]), \tag{2.32}$$

where $J[r]$ is the Jacobian of the coordinate transformation. (Actually, in this case $J[r] = r$.) Equation (2.32) is however wrong! The reason for this can qualitatively be given as follows:

If we discretize the time derivatives $\dot{x}^2+\dot{y}^2$, we have

$$\begin{aligned}
\dot{x}^2+\dot{y}^2 &\simeq [(x_i-x_{i-1})^2+(y_i-y_{i-1})^2](\Delta t)^2 \\
&= [(r_i-r_{i-1})^2+2r_ir_{i-1}(1-\cos(\phi_i-\phi_{i-1}))]/(\Delta t)^2 \\
&= [(r_i-r_{i-1})^2+r_ir_{i-1}(\phi_i-\phi_{i-1})^2 \\
&\quad -\frac{1}{12}r_i^2(\phi_i-\phi_{i-1})^4+\cdots]/(\Delta t)^2.
\end{aligned} \tag{2.33}$$

If the paths were differentiable the third term would vanish as $\Delta t \to 0$. But, as we know, this is not the case with Wiener integrals. We deduce from equation (2.33) that terms of order r^2/t^2 contribute to the integral in question. Note that if we also consider the summation, we have the general result that significant terms in the discretization procedure are at least of the order $(\Delta x)^2/\Delta t$. This is similar to the property $(dX)^2 \sim dt$ we observed in stochastic integration, and due to the same reason: the nondifferentiability of the paths of Brownian motion.

Because of the very same reason we must also specify what we mean by x when we discretize the Wiener integral. This choice has the same effects as in stochastic integration: we can choose the point to be x_i, or in a general case, $x_i+\lambda(x_{i-1}-x_i)$, where $\lambda \in [0,1]$. For $\lambda = 0$ we have the above case – the postpoint prescription. For $\lambda = \frac{1}{2}$ and $\lambda = 1$ we get the midpoint and prepoint prescription, respectively.

To give an example of how the value of an functional integral depends on the discretization rule used we calculate the Wiener integral

$$I = \int \mathcal{D}x\, e^{-\int_{t_0}^{t} dt\, \mathcal{L}},$$

where

$$\mathcal{L} = \frac{1}{2}\dot{x}^2 + \dot{x}x + \frac{1}{2}x^2$$

and the endpoints on the interval $[t_0, t]$ are fixed. We begin by discretizing the integral in a general scheme i.e with an arbitrary $\lambda \in [0,1]$. This yields

$$I_N = \int \prod^{N} dx_i \prod^{N+1} \frac{1}{\sqrt{2\pi\varepsilon}} \exp(-\sum^{N+1} \varepsilon \Big[\frac{1}{2}\frac{(\Delta x_i)^2}{\varepsilon^2} + ((1-\lambda)x_{i-1} + \lambda x_i)\frac{\Delta x_i}{\varepsilon} + \frac{1}{2}((1-\lambda)x_{i-1}^2 + \lambda x_i^2)\Big]). \tag{2.34}$$

The discretized form (2.34) can be worked out by the Gel'fand-Yaglom method [G&Y60]. This includes calculating the determinant of an $N \times N$ matrix. In the end we can take the limit $N \to \infty$ and we have the result

$$I = [\pi e^{2\lambda(t-t_0)}(1 - e^{-2(t-t_0)})]^{-1/2} e^{-(x - x_0 e^{-(t-t_0)})^2/(1-e^{-2(t-t_0)})}. \tag{2.35}$$

Equation (2.35) explicitely shows that the value of the functional integral depends on λ; the way we discretize the integral. Note that (2.35) is now the Green's function of the partial differential equation

$$\partial_t \phi(x,t) = [\frac{1}{2}\partial_x^2 + \partial_x x - \lambda]\phi(x,t),$$

where ∂_x means $\frac{\partial}{\partial x}$.

(This example is due to Langouche et al [LRT82].)

The form of the Wiener integral (2.30) is almost identical to the well-known Feynman path integral of quantum mechanics. Even though these two integrals are so close to each other – in fact they are connected by the transformation $t \to it$ – their mathematical properties are quite different. In the next chapter we will review the properties of the Feynman path integral and its similarities and differences with the Wiener integral.

Chapter 3

Feynman Path Integrals

The concept of applying path integration to quantum mechanics dates back to the original paper of Feynman [Fey48], based on his unpublished doctoral dissertation of 1942. Actually, Dirac had prompted the usage of classical Lagrangian function in the description of time evolution of quantum mechanical systems already in the thirties [Dir33]. Nevertheless, it was Feynman who in his famous paper emerged with the idea of expressing quantum mechanics in terms of functional integration. Feynman's paper has had an enormous impact on almost every branch of physics; even today, his paper is referenced in many papers.

The original idea was that the transition amplitude (roughly equivalent to the transition probability of section 2.4) of a quantum mechanical particle could be written as the sum over the possible histories – or paths – of the particle. In mathematical terms,

$$K(x_1,t_1,x_0,t_0) \sim \sum_{\text{all paths}} e^{i\mathcal{S}[x(t)]}, \tag{3.1}$$

where $\mathcal{S}$ is the action integral $\int_{t_0}^{t_1} dt\, \mathcal{L}(x(t),\dot{x}(t),t)$ of classical mechanics. Later on, it is conceptually very easy to regard the above sum as an integral over the space of paths.

Notice that (3.1) is almost the same as equation (2.30) with $f(x) \equiv 1$. The small but important difference is having i instead of -1 multiplying the action. This little difference opens up an enormous gap between the two integrals when one analyzes them mathematically.

We will first present Feynman's definition of path integrals. Then we discuss how this concept can be further developed and what kind of mathematical problems one encounters if one tries to define Feynman path integration in a mathematically solid way. Our basic construction of the Feynman path integral follows closely the lines of Feynman's own book [F&H65]. In this work we will only discuss nonrelativistic path integrals to keep in mind the connection with

the original Feynman path integrals. Path integration has, of course, been developed for relativistic quantum theory as well and it is an indispensable tool in quantum field theory.

3.1 Definition

Feynman's definition of path integrals rests on the notion of *probability amplitude*, ϕ. Probability amplitudes are usually complex numbers; the probability corresponding to a certain probability amplitude is given by the absolute square, $P = |\phi|^2$, of the amplitude. The basic property of probability amplitudes is that whenever there are two possible ways for an event to happen, the probability for such event is *not* the sum of the probabilities, $P_1 + P_2$ but the absolute square of the sum of the probability amplitudes, $P = |\phi_1 + \phi_2|^2$. This is due to quantum mechanical phenomena, for example the motion of electrons through a double slit which shows an interference pattern similar to the behaviour of waves in classical mechanics. Probability amplitudes are thus frequently called *wave functions*.

We begin by assigning a probability amplitude to the event that a particle travels from point x_0 to point x in configuration space. Let us first assume that there are slits on the way of the particle, so that there are only finite number of paths the particle can actually move along. For each such path we assign a separate probability amplitude. As stated above, the total probability amplitude is then the sum over the amplitudes of all these paths. This setting is similar to the equation (2.23), where we demanded that the particle must move through a given set of points.

As with Wiener integrals, it is natural to allow x_j to take arbitrary values at each fixed t_j. Integrating over the variables x_j gives us the probability amplitude for the particle to move along a free path. It is quite natural to expect this to extend to possible all paths from x_0 to x by letting the number of intermediate points approach infinity.

But what exactly is the probability amplitude corresponding to a single path? Feynman tells us that the phase of the amplitude is proportional to the classical action. This is where he used the proposal of Dirac to connect concepts of classical mechanics with quantum mechanics. Having done that, we arrive at the expression for the probability amplitude given by equation (3.1).

What we have is still just the limit $N \to \infty$ of the product of integrals

$$K(x_1, t_1, x_0, t_0)_N = A^N \int_{\mathbb{R}\times\cdots\times\mathbb{R}} \prod_{j=1}^{N-1} dx_j \; e^{i\mathcal{S}[\{x_1, x_2, \ldots, x_{N-1}\}]}, \tag{3.2}$$

where $A = \frac{1}{\sqrt{2\pi i(t_1 - t_0)}}$ is a normalization factor. It is tempting to assume that (3.2) has as its limit an integral corresponding to the Wiener integral. Feynman

thus wrote the above product *formally* as an functional integral:

$$K(x_1,t_1,x_0,t_0) = \int \mathcal{D}_F x\, e^{i\mathcal{S}[x(t)]}, \tag{3.3}$$

where $\mathcal{D}_F x$ is the Feynman "measure" similar to the Wiener measure. More precisely,

$$\mathcal{D}_F x \simeq \lim_{N\to\infty} \prod_{j=1}^{N-1} \frac{dx_j}{\sqrt{2\pi i\varepsilon}}, \tag{3.4}$$

where $\varepsilon = (t_1 - t_0)/N$. Then he named (3.3) a *path integral*. It must be emphasized that notation like equation (3.3) is meant to be used as a generalizing notation of the concept of path integration; it is not meant to be a rigorous formulation of path integration. Unfortunately it is exactly the equation (3.3) that many take as the definition of Feynman path integrals.

Notice that if we take $t \to -it$ and substitute $\mathcal{D}_F x$ for $\mathcal{D}_W x$ we have the Wiener integral (2.26). This fact that the two integrals have almost the same mathematical structure is important; explicit solutions of the integrals of one type can usually be applied immediately to the other type. We can also interpret the Feynman path integral as a Wiener integral of pure imaginary variation. Because the actual calculation of Wiener integrals uses nearly the same techniques as Feynman path integrals we have a large number of calculational methods at our hand when we need to find the value of a Feynman path integral.

It must stressed that there is only a small number of physical systems for which path integrals can be evaluated explicitly. Fortunately, one of the benefits of path integrals is that they allow a intuitive and effective way of solving the integrals numerically, usually by Monte Carlo simulation. The number of analytically solvable path integrals has grown rapidly in the last decade and now almost every quantum mechanical system solvable by Schrödinger equation can also be solved by path integral methods. A review of analytical calculational methods of path integrals is given by Grosche and Steiner [G&S95].

This procedure of taking the mathematically sound Wiener integrals as the starting point and then *analytically continue* them to imaginary time is quite common. One of the early developments of this idea was by Nelson [Nel64]. Analytical continuation especially useful in constructive field theory, where the Euclidean field theory (with imaginary time) has been studied extensively by this method. In some cases the results can then be transferred to quantum field theory with normal time. To what extent this analytical continuation is possible has been studied by Osterwalder and Schrader ([O&S73],[O&S75]), for example.

Like Brownian motion, the paths of quantum mechanics are also fractals. Abbott and Wise [A&W81] show this explicitely. They also point out that the

transition from classical mechanics to quantum mechanics can be seen in the fractal nature of the paths of particles; the paths of quantum mechanics have Hausdorff dimension of 2, while for classical paths it is 1. This reflects the uncertainity principle of the quantum world – the path is effectively two-dimensional because of the randomness of the movement.

Since the original definition of Feynman path integrals has its weak points and is not mathematically rigorous a lot of research has been done to find an alternative way to define the integral. Several different attempts have been made; some of these will be exposed and discussed in section 3.3. But first we will list the problems of the above definition. Note that the critique is directed at the above original definition of the Feynman path integrals.

3.2 Problems of definition

Even though the definition (3.3) seems to be complete from a physical point of view, mathematically it is on a very shallow ground as Feynman himself admits [F&H65]. There are several mathematical reasons why (3.3) is not a functional integral in a mathematical sense. In fact, the very construction of the integral is built mainly on physical concepts, with hardly any mathematics involved. Comparing this with the definition of Brownian motion and the Wiener integral, we find that there is a large gap between the two integrals when it comes to mathematical rigour.

A fine review of the following mathematical problems and an account of the work done on them is given by Tarski [Tar74]. The problems he mentions will be discussed in this work. We shall also make additional remarks concerning the reasons for mathematical problems in Feynman path integrals.

First of all, the introduction of the classical action in the determination of the probability amplitude may seem an arbitrary choice. But, if we consider quantum mechanics in terms of operators, we find that the Hamiltonian operator (and thus the classical Hamiltonian function and Lagrangian function) is the generator of time-evolution for states. Also, if we think of the Feynman path integral as the Green's function of the Schrödinger equation we find that the inclusion of the classical action in the exponential term follows from Feynman-Kac formula in the same way as with Wiener integrals. So, at least from a physical point of view the form of Feynman path integrals is not problematic.

Another argument against the validity of (3.3) is that the exponential term includes the imaginary unit i and thus the integral does not converge in the usual sense. The usual "solution" is to include a small imaginary component in the exponential term – usually it is assumed that $\hbar$ is partly imaginary, say $\hbar + i\varepsilon$. This makes the integral to converge as a normal Gaussian integral. After calculating

the integral we take the limit $\varepsilon \to 0$. The use of this so-called "$+i\varepsilon$"-prescription is actually prompted by causality arguments when we interpret the path integral as a Green's function of the Schrödinger equation.

It is also possible to carry out the "$+i\varepsilon$"-prescription by adding a small imaginary part to the mass of the particle, which is not explicitely shown in our formulae. This has the same convergence-yielding effect but in addition has other effects on the integral. Hence the choice of which parameter of the path integral is subject to analytical continuation depends largely on the physical situation in question and what one wishes to calculate with the path integral.

The oscillation of the exponential term can also be a used in a positive way to show an illuminating detail: if we write (3.3) with $\hbar$ correctly inserted, we have

$$K(x_1,t_1,x_0,t_0) = \int \mathcal{D}_F x \; e^{\frac{i}{\hbar}\mathcal{S}[x(t)]}. \tag{3.5}$$

Notice that for $\mathcal{S} \gg \hbar$ the exponential term oscillates strongly for all paths except in the vicinity of the classical path which minimizes the action functional. For all other values of the action the phases are approximately opposite and thus cancel each other out. In effect, the only paths that affect the value of the integral correspond to the classical path. This example gives a very intuitive and enlightening view to the fact that if $\hbar \to 0$ we should get classical mechanics from quantum mechanics. We also see that for $\mathcal{S} \sim \hbar$ quantum phenomena cannot be neglected as we must consider other paths in addition to the classical path when calculating the integral (3.5).

If we were to understand the equation (3.3) as a functional integral, there are still several unclear questions, such as what are the integrable functions and what exactly is the underlying space of paths. Feynman did not originally answer either of these questions. Only experience has taught us which functions are integrable – and thus which potentials can be handled with the path integral formalism. This connection can be seen as follows: we assume that the Lagrangian is of the type $L = T - V$. Usually the kinetic term is included in the definition of the path integral and so the integrated functional is $\exp(i\int dt\; V[q(t)])$. (This is the Feynman-Ito formula, see Theorem 3.1.) As we shall see in section 3.3, all the alternative definitions do declare what functionals are integrable.

3.2.1 Probability amplitude

The basic source of trouble in Feynman path integrals is the concept of probability amplitude itself; for a free particle i.e. a system without a potential it is

$$K(x,t,0,0) = \int \mathcal{D}_F x \; \exp(i\int_0^t ds \frac{1}{2}\dot{x}^2), \tag{3.6}$$

assuming that the particle starts at $t = 0$ from the origin. The above path integral can be evaluated exactly and the result is

$$K(x,t,0,0) = \frac{1}{\sqrt{2\pi i t}} e^{ix^2/2t}. \tag{3.7}$$

Although the amplitude has some of the properies of a transition probability – for example, the Einstein-Smoluchowski-Chapman-Kolmogorov equation still holds for successive transitions – it misses the most important feature of the transition probability of Brownian motion: it is not a measure.

Let us further scrutinize the equation (3.7). As can be seen, the fact that it is not a real positive measure is due to the imaginary factor i present in the exponential term. It can be easily checked that (3.7) does not fulfill the properties of a probability measure (compare with definition 2.1). One can generalize the theory of measures to complex measures $\mu : \Omega \to \mathbb{C}$ but (3.7) is not even a complex measure. This is due to the oscillation of the exponential term; although we cannot require the unit measure $\mu(\Omega) = 1$, the property $\mu(A \cup B) = \mu(A) + \mu(B)$ should hold for $A, B \subset \Omega$, $A \cap B = \emptyset$. Because of the oscillation this is not true for the Feynman measure.

The genuine probability is the absolute square of the probability amplitude. For (3.7) it is

$$P(x,t) = (K(x,t,0,0))^2 = \frac{1}{2\pi t}. \tag{3.8}$$

Even more trouble follows; if we integrate over the endpoint to calculate the total probability, the result is infinite and not 1, as it should. In fact this is why the transition probability (3.6) cannot be a measure – it does not converge [Cam60]. The above expression must be a relative probability. This also prompts us to use path integration in the context of time evolution of wave functions.

An interesting interpretation of the fact that the probability (3.8) does not converge for a free particle is as follows. Recall that a free particle, with momentum p, is represented by a wave function $\psi(x) = e^{ipx}$. This function is not square integrable; $\int dx|\psi(x)|^2 = \infty$. Thus we usually apply some sort of normalization, for example by restricting the particle to a box with finite dimensions. Then we can carry out calculations and in the end we let the dimensions of the box go to infinity. The same applies for the propagator; since the free particle is not localized, we should not expect its probability to converge.

Since even the short-time probability amplitudes fail to be measures (in any reasonable sense) it is impossible to proof that there exists a limit of these measures defined on discrete intervals which is equal to a measure defined on the space of paths. The lack of a well-defined measure prohibits us from developing rigorous integration theory for Feynman path integrals. The Feynman path

integral is only a linear functional and not a functional integral. Equations such as (3.3) must therefore be understood as a formal way of writing down the limit of the product of integrals. The reason why path integrals are so popular is that even though they are not really integrals they behave exactly like genuine integrals under linear coordinate transformations.

The formal definition has its benefits, though. Intuitively it is clear what is meant by such an equation and from physicist's point of view it is mathematically correct enough to be used as a tool in quantum mechanics. The given notation is also independent of how we are actually define the integral. If we ever find a solid definition for the path integral, we need not change our notation.

3.2.2 Hamiltonian formalism

Another shortcoming of the original definition is that it is based on concepts derived from classical mechanics rather than quantum mechanics. In quantum mechanics based on operators, the Hamiltonian operator is the basic ingredient, not the Lagrangian of Feynman's definition. It is possible to define path integrals in terms of the Hamiltonian, and many authors actually take this as their starting point. In can be argued that this approach is more general and better suited for quantum mechanics in which the Hamiltonian operator plays a major role. These two separate formulations of quantum mechanics are connected by the time evolution operator, $\exp(it\widehat{H})$, where $\widehat{H} := H(\widehat{x},\widehat{p},t)$ is the Hamiltonian operator.

The proper discussion of the use of $\widehat{H}$ in the time evolution operator involves determining the order of the operators $\widehat{x}$ and $\widehat{p}$. This is important since they do not commute. We postpone this discussion to section 3.4.1 where we consider the problem in detail. For the time being we assume that operator $\widehat{H}$ is of the form $\widehat{H} = T(\widehat{p}) + V(\widehat{x})$ and then no problem of operator ordering exists.

Using this, we can write the transition amplitude from state $|x_0\rangle$ to state $|x\rangle$ as

$$K(x_0,x,t) = \langle x_0|e^{it\widehat{H}}|x\rangle. \tag{3.9}$$

$K(x_0,x,t)$ is actually, as we shall see, equivalent to the transition amplitude of Feynman's path integral. They both are Green's functions of the Schrödinger equation. And of course, $K(x_0,x,t)$ of the equation (3.9) can be expressed as a path integral. To see this, we first note that for time evolution operators $\widehat{U}(t,t_0) := exp(-i(t-t_0)\widehat{H})$ we have the rule

$$\widehat{U}(t,t_0) = \widehat{U}(t,t_1)\widehat{U}(t_1,t_0), \tag{3.10}$$

where $t_0 \leq t_1 \leq t$. (this is in fact equal to the ESKC-equation (2.22))

Let us now divide the interval $[t_0,t]$ into N equal subintervals. We can then write down the transition amplitude from x_0 to x as

$$K(x_0,t_0,x,t) = \langle x|\widehat{U}(t,t_{N-1})\widehat{U}(t_{N-1},t_{N-2})\cdots\widehat{U}(t_1,t_0)|x_0\rangle. \tag{3.11}$$

If we insert the unit operators

$$\widehat{1} := \int_{-\infty}^{\infty} dx_n |x_n\rangle\langle x_n|, \quad n = 1,\ldots,N-1$$

in suitable positions in equation (3.11) we have the product of $N-1$ integrals

$$K(x_0,t_0,x,t) = \int \prod_{j=1}^{N-1} dx_j \prod_{j=1}^{N} K(x_j,t_j,x_{j-1},t_{j-1}). \tag{3.12}$$

The entity $K(x_j,t_j,x_{j-1},t_{j-1})$ is called the *short time propagator*. If we take $t_j - t_{j-1} = \varepsilon = (t-t_0)/N$ it can be written in the form

$$K(x_j,t_j,x_{j-1},t_{j-1}) = \langle x_j|e^{-i\varepsilon\widehat{H}}|x_{j-1}\rangle$$

Now, as we have assumed, the Hamiltonian operator $\widehat{H}$ can be written in the form $\widehat{H} = T(\widehat{p}) + V(\widehat{x})$. We can calculate the short time propagator with the help of the Baker-Hausdorff formula

$$e^{-i\varepsilon\widehat{H}} = e^{-i\varepsilon\widehat{V}} e^{-i\varepsilon\widehat{T}} e^{-\varepsilon^2\widehat{X}}, \tag{3.13}$$

where $\widehat{X}$ is an operator which depends on the the commutators $[\widehat{V},\widehat{T}]$. If $\widehat{V}$ and $\widehat{T}$ commute then $\widehat{X} = 0$. For non-commuting operators we can usually neglect the terms which are of order ε^2. The short-time propagator is then

$$K(x_j,t_j,x_{j-1},t_{j-1}) \approx \int \frac{dp_j}{2\pi} exp\{ip_j(x_j - x_{j-1}) - i\varepsilon H(p_j,x_j)\},$$

where we have replaced the Hamiltonian operator $\widehat{H}$ with the Hamiltonian function H. This procedure of finding the corresponding classical function will be fully described in section 3.4.1. The approximation sign is due to the fact that $\widehat{x}$ and $\widehat{p}$ do not commute. The above equation is true to the order of ε^2.

The transition amplitude for the interval $[t_0,t]$ can then be written as

$$K(x_0,t_0,x,t) \approx \int \prod_{j=1}^{N-1} dx_j \prod_{j=1}^{N} \frac{dp_j}{2\pi} \exp(i\mathcal{S}^N), \tag{3.14}$$

where $\mathcal{S}^N$ is the sum

$$\mathcal{S}^N = \sum_{j=1}^{N} [p_j(x_j - x_{j-1}) - \varepsilon H(p_j, x_j)]. \tag{3.15}$$

For a reasonably smooth Hamiltonian function H the right hand side of equation (3.14) converges and the approximate sign can be replaced by an equality when we take the limit $N \to \infty, \varepsilon \to 0$. Note that in this limit the sum (3.15) tends toward the classical canonical action for a path $[x(t), p(t)]$ in phase space. Like the original Feynman path integral, the limit of equation (3.14) can be written formally as

$$K(x_0, t_0, x, t) = \int \mathcal{D}'x \, \frac{\mathcal{D}p}{2\pi} e^{i\mathcal{S}[p,x]}. \tag{3.16}$$

The prime in the above equation is included because the endpoints x_0 and x_t are fixed for all paths and thus there is one less integration over coordinates in equation (3.14) than there is over canonical momenta.

The equation (3.16) can be understood as the generalization of the configuration space path integral, although they are derived from quite different basic assumptions. As a matter fact, for Hamiltonian functions with only quadratic dependence on p we can readily integrate over the momenta in equation (3.12). The result is the configuration space path integral as defined by Feynman. It must be stressed that they all are integrals only formally – there is no measure defined on an infinite-dimensional phase space and therefore no true integration theory on the paths of phase space.

If we take $t \to it$ in integral (3.16) we see that the resulting integral is actually similar to one we find in statistical physics – the partition function. The use of path integrals in statistical physics is a broad subject itself and outside the scope of this work. Therefore we shall not delve deeper into it. Let us just recall that path integration is a very important mathematical tool in statistical physics.

However, the procedure of defining path integrals beginning with operator approach to quantum mechanics takes us even further away from the mathematically rigorous theory of Brownian motion. The paths in the phase space are not even continous – this can be seen by considering the configuration space integral and recalling that the paths are nondifferentiable; thus the momenta is discontinous almost everywhere. The functional integral defined through operators and Hamiltonian formalism is even more clearly only a formal expression meaning a limit of a product of integrals than in the case of configuration space path integral.

If we use operator formalism to define the quantum mechanical path integral we may fail to appreciate the fact that the way Feynman originally defined his

path integrals is completely independent of the operator formulation of quantum mechanics.

As with the original Feynman path integral, a casual physicist may use the formal equation (3.16) as the definition of the phase space path integral and discretize it to find explicit solutions to the integral. However, from the discussion above we know that the definition is actually built on this discretized form of the transition amplitude.

3.3 Alternative definitions

Considering the multitude of mathematical problems connected to the original definition of path integration, as shown in the previous section, it is no wonder that a lot of work has been done to build a proper foundation for path integration. Unfortunately, as with Wiener integrals, there are many ways to look at the Feynman path integral. One's viewpoint strongly affects what one takes as the starting point of the definition. In the following we discuss some of the best attempts at defining the Feynman path integral in mathematically proper terms.

We must recall that the attempts are somewhat indirect; the structure of the integral is such that it does not allow a direct approach since we cannot define a normal measure on a function of spaces that coincides with the ordinary definition of the Feynman path integral.

The constructs used to alternatively define the path integral can be divided into two classes: the first deals directly with the space of paths and tries to make up for the lack of a proper measure by the means of some abstract mathematical construct. The other type uses the discretized version of the path integral as its basis.

The following definitions are based on slightly different assumptions. Hence the class of integrable functionals varies, as mentioned in section 3.2. This class of integrable functionals is usually defined in terms of the potential part V of the Lagrangian. Note that this does not necessarily imply anything about the class of physical potentials which can be handled with the specific definition; it merely tells that the certain functionals fulfill the integrability conditions of the definition.

3.3.1 First constructs

The first notable attempt to redefine the Feynman path integral was made by Cameron in 1960 [Cam60]. He observed that usual measure-theoretic models cannot be used in connection with Feynman path integrals even if the exponent of $\exp(i\mathcal{S})$ is modified to have a real part. He then used a technique similar to

Wiener integrals to redefine path integrals; namely, he divided the interval $[t_0, t]$ into n subintervals with

$$t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t.$$

The paths are the assumed to be polygonal. Functionals $\exp(i\mathcal{S})$ are the calculated along these polygonal paths and finally when one passes to the limit $\max(t_j - t_{j-1}) \to 0$ one arrives at a construct one can call a path integral. Note that one needs to apply the modification scheme $m \to m + i\delta$ in order to calculate the integrals over the discretized coordinates x_i.

This approach is somewhat analogous to the Hamiltonian formalism in section 3.2.2 and defines the path integral as the limit $N \to \infty$ of integral in finite number of dimensions. The problem is that, as Cameron himself admits, there does not exist a measure nor an integral on the space of continous functions that would coincide with the finite-dimensional integral. Therefore the path integral defined this way is only an heuristic derivation and is only a formal notation similar to the original definition by Feynman.

The above definition can be also achieved by using the Trotter formula

$$\exp[t(A+B)] = \lim_{n\to\infty} [\exp(tA/n)\exp(tB/n)]^n.$$

We take $A = iH_0$, where H_0 is the free Hamiltonian ($\frac{1}{2}p^2$) and $B = -iV$. This was first done by Nelson [Nel64].

Another approach – which we are already familiar with – is also due to Nelson [Nel64] who defined Feynman path integrals by analytical continuation of Wiener integrals. As we know this is doomed to fail because when we make the transformation $m \to im$ in Wiener integrals the variance becomes purely imaginary and the Wiener measure becomes a complex measure with infinite total variation. This is exactly the case of Cameron explained above; without a proper measure we cannot have a proper integral.

Both of the above alternative definitions, however, have the same shortcomings. First of all, they are both indirect in their approach to the path integral i.e. neither actually manages to directly write the path integral as a true integral of the form

$$I = \int \mathcal{D}x\, e^{i\mathcal{S}[x]}, \tag{3.17}$$

where $\mathcal{S}[x]$ is the classical action. The above integral is what we conceptually mean when we speak of functional integral in quantum mechanics and therefore the definition we use should include an integral of the type (3.17). We also recall that the definitions are quite restrictive with respect to the potential term V in the Lagrangian – V must satisfy strong conditions such as analyticity.

Another weakness of the above definitions is that they do not directly support the most important intuitive feature of Feynman path integrals, namely the transition to classical mechanics as $\hbar \to 0$.

A different approach to the problem was taken by Ito [Ito66], who suggested the definition of the path integral through measures $d\mu_{T,\alpha}$ on a Hilbert space $\mathcal{H}$ with inner product $< x,y >$. $d\mu_{T,\alpha}$ is defined by the covariance operator T (symmetric, strictly positive) and the mean vector α. We denote

$$I_{T,\alpha}(f) = \frac{\int d\mu_{T,\alpha}(x)\, e^{i<x,x>} f(x)}{\int d\mu_{T,0}(x)\, e^{i<x,x>}}.$$

A path integral on such space is then defined as

$$I = \int \mathcal{D}x\, e^{i<x,x>} f(x) = \lim_{T\to\infty} I_{T,\alpha}(f). \tag{3.18}$$

The limit $T \to \infty$ must be taken suitably and must be independent of the vector α. Ito proved the convergence of the above integrals for a very limited class of functionals f with respect to potentials V. Ito's proposal is too limited to be a general definition of path integration: it only works for potentials which are Fourier transforms of bounded complex measures and for potentials of the polynomial type

$$V(x) = \sum_{i=1}^{2} c_i x^i,$$

with $c_2 > 0$.

3.3.2 Pseudomeasures

The first formulation of Feynman path integrals that has a solid mathematical foundation was made by C. DeWitt-Morette [DeW72]. The definition she used has been later reviewed and refined in [DMN79] and in [C&D95]. It also paved the way for further alternative definitions based on slightly different assumptions (see the following sections) which nevertheless used the same way of indirectly defining the measure on a function space

The definition is built on the concept of *pseudomeasures* or *prodistributions*, prompted by the work of Bourbaki on *promeasures*. (see, for example, [Bou69]). The theory of promeasures generalizes the theory of measure to spaces which are not locally compact, which is the case in most spaces of functions. Pseudomeasures are defined as the Fourier transforms of Gaussian measures. One should be aware that even though they are called "measures" they are not usual set-theoretic measures. Slightly altered, this formalism can also be applied to the

Wiener integrals at once, which makes this approach to functional integration very general.

Let us now consider the space of continous paths $\mathcal{C}$ from $T :=]t_a, t_b[\to \mathbb{R}$, with the norm $||q|| = \sup_{t\in T} |q(t)|$. The space $\mathcal{C}$ can be understood as a topological space which is Hausdorff and locally convex. We also assume that $\mathcal{C}$ is a linear space in the sense that for two paths $q, q' \in \mathcal{C}$ also the sum of the paths $q + q'$ is also in $\mathcal{C}$. This is achieved by demanding that $q(t_b) = 0$ for all $q \in \mathcal{C}$.

The space $\mathcal{C}$ has a dual space $\mathcal{M}$, which is the space of bounded measures defined on T. The space of paths and its dual are connected by the equation

$$< \mu, q >= \int_T q(t) d\mu(t) \quad \in \mathbb{C}. \tag{3.19}$$

(Note that $\mathcal{M}$ includes complex measures.)

If X is a topological space and X' is its dual space, the Fourier transformation is defined with the help of the above equation as

$$\mathcal{F}\{\lambda(x')\} := \int_X e^{-i<x',x>} d\lambda(x), \tag{3.20}$$

where λ is a measure on X'. It should be noted that the above equation defines a infinite-dimensional Fourier transform. Now we can define the bounded Gaussian measure w on $\mathcal{C}$ as

$$\mathcal{F}\{w\} = \exp(-\frac{i}{2} W), \tag{3.21}$$

where W is the covariance of the measure μ,

$$W(\mu) = \int_T \int_T \inf(t,t') d\mu(t) d\mu(t'). \tag{3.22}$$

(Compare this with the characteristic function of the transition probability of Brownian motion, equation (2.3).) The choice of $\inf(t,t')$ is made to connect this definition to the ordinary Feynman path integrals. Note that W is a positive quadratic form on $\mathcal{M}$. In the general case of covariance function $K(t,t')$ we have

$$\mathcal{L}(\mu) = \int \int K(t,t') d\mu(t) d\mu(t') \tag{3.23}$$

and w is defined by the equation

$$\mathcal{F}\{w\} = e^{-\frac{i}{2}\mathcal{L}(\mu)}. \tag{3.24}$$

Having defined the measure w we can proceed to define the path integral on the space $\mathcal{C}$ in the form ($m = 1$)

$$K(b,a) := \int_{\mathcal{C}} e^{i\mathcal{S}_{\text{int}}(q)} dw(q). \tag{3.25}$$

We assume that the Lagrangian is of the form $L = T - V$ so that $\mathcal{S}_{\text{int}} = \int dt\, V(q(t))$. The integration measure w is the measure defined by the equation (3.21). Note that we have not used discretization in any step of the above definition. The above integral may thus be a true functional integral on the space of paths.

The main point of this definition is that it coincides with Feynman's original definition of the path integral i.e.

$$\int_{\mathcal{C}} dw(q) = \int_{\mathbb{R}} \exp\Big(i \sum^{n} \frac{(\Delta q)^2}{2\varepsilon_j}\Big) \prod^{n} \frac{dq_j}{\sqrt{2\pi i \varepsilon_j}}, \tag{3.26}$$

with $\varepsilon_j := t_j - t_{j-1}$ and $(i)^{-1/2} = \exp(-i\pi/4)$. The right hand side of equation (3.26) can be recognized as the original path integral with n discretization points; we denote it as I_n. DeWitt-Morette proves in her paper [DeW72] that the "measure" $dw(q)$ is equal to the product of "measure" of transition amplitudes. This is completely analogous to the Wiener measure which is also equal to the product of measures on a sequence of transition probabilities, cylinder measures.

One can easily see the similarities in the mathematical structures; we have a measure defined on a space of paths (or functions) which coincides with the product of a finite number of measures on cylinders. This equality has no value in the context of the definition but is very important if we wish to actually calculate a functional integral. These similarities are well-documented in the reference [C&D95], where Cartier and DeWitt-Morette develop the theory of functional integration from the basis of the notion of pseudomeasures.

Unfortunately, as mentioned before, the transition amplitude is not a measure in the usual sense. This is true also for the "measure" w defined by the equation (3.21). One must therefore make a generalization of integration theory to spaces which are not locally compact, for example spaces of paths, and then use other arguments show that this definition is reasonable and meaningful. These arguments include that one should understand the measure as a bounded distribution of rank zero and not as a measure, since it is an unbounded measure. This interpretation makes the equation (3.26) sensible even though we are not dealing with true measures. The point is that even though w is a poor measure it is a good distribution.

To discuss the distributional properties of the pseudomeasure w we assume a functional $\phi \in \mathcal{C}^{\infty}(\mathbb{R})$, which is the space of smooth functions on $\mathbb{R}$. We will use ϕ as a test function. Then the inner product

$$< w, \phi >= \int \phi(q) dw(q)$$

is always defined. Furthermore, w is in the space of multiplication operators O_M. If $T \in O_M$ then

$$< wT, \phi >=< T, w\phi >$$

is defined for all T. Pseudomeasure w belongs also to the space of convolution operators. These distributional properties are actually the ones that prompt us to use Fourier transformation to define the desired measure on the space of paths.

When we are to explicitely calculate functional integrals, what we need to do is to calculate the inner product according to the equation (3.19). For example,

$$\int_C dw(q) =< w,1 >; \quad \int_C e^{iS(q)} dw(q) =< w, e^{iS} > . \tag{3.27}$$

The general case is then calculated as ($\overline{\mathcal{F}}$ means the inverse Fourier transform)

$$< w,\phi >=< \overline{\mathcal{F}}\mathcal{F}w,\phi >=< \mathcal{F}w,\overline{\mathcal{F}}\phi >=< e^{-iW/2},\overline{\mathcal{F}}\phi > . \tag{3.28}$$

The final term is calculated like the equation (3.19). An important example of the above integrals is

$$< w,1 >=< e^{-iW/2},\overline{\mathcal{F}}1 >=< e^{-iW/2},\delta(0) >= 1. \tag{3.29}$$

Although it may seem that the definition is rather abstract and that it is hard to derive explicit results using it, this is not the case. The property (3.26) connects these integrals readily to ordinary Feynman path integrals on $\mathbb{R}^n$. This connection is made clearer in the paper [DeW74], where DeWitt-Morette clarifies the definition and uses the technique to calculate some simple path integrals. She also shows that the definition allows one to make semiclassical approximations of the functional integral with ease, which is a good feature since the possibility to use classical approximation is one of important attribute of Feynman path integrals.

The procedure can be generalized to phase space path integrals. This work has been done by Mizrahi ([Miz76],[Miz78]). He also develops computational techniques based on this definition of Feynman path integrals.

There are limitations to the pseudomeasure approach. First of all, the class of potentials V for which this formulation is valid was not originally defined clearly. In [DeW74] DeWitt-Morette shows that the formalism works for linear continous potentials and for non-linear potentials $V(q)$ which satisfy the constraint

$$\int_{\mathbb{R}} du\, V(u) e^{iu^2} < \infty.$$

This class includes, for example, all polynomial potentials.

The validity of several other arguments is somewhat ambiguous as well. Furthermore, only paths belonging to the function space $L^{2,1}$ are considered. By $L^{2,1}$ we mean the space of functions whose derivative is square integrable, $\int dt|\dot{q}|^2 < \infty$. From the considerations of chapter 2 it is clear that the definition

of path integrals should also allow integration over continous functions which are not differentiable.

The definition via pseudomeasures may even be critized for being too general – in most cases we make further assumptions on the space of paths and hence removing some of the complications of the more general case. In spite of these weaknesses DeWitt-Morette's ideas have been very important as the inspiration and basis for further attempts to redefine the Feynman path integral as we will see in the following sections.

3.3.3 Fresnel integrals

Another attempt to rewrite the definition of Feynman path integrals is due to Albeverio and Høegh-Krohn [A&H76]. They also take an indirect way to define the path integral, this time by a construct called the *Fresnel integral* (after similar integrals found in optics). Their approach is built around Hilbert space formalism rather than topological concepts. They define their "measure" on a Hilbert space of continous functions by a Fourier transform. In the following we shall briefly study the main features of this definition which is based on the concept of an *oscillatory integral* i.e. a functional integral of the form

$$\int \mathcal{D}x\, e^{iA}.$$

We begin by considering the Hilbert space of $\mathcal{H} = \mathbb{R}^n$ with inner product denoted by (x,y). The norm is then $|x|^2 \equiv (x,x)$. Let us first write $N = (2\pi i)^{\frac{n}{2}}$. With this notation we write the Fourier transform of the function $\phi(x)$:

$$\mathcal{F}\{\phi\}(x) = \int dy\, e^{i(x,y)}\phi(y). \tag{3.30}$$

We also note that

$$\frac{1}{N}\int dx\, e^{\frac{i}{2}|x|^2}\mathcal{F}\{\phi\}(x) = \int dx\, e^{-\frac{i}{2}|x|^2}\phi(x) \tag{3.31}$$

Let $f(x)$ now be the Fourier transform of a bounded complex measure μ (thus $||\mu|| := \int |d\mu| < \infty$) on $\mathcal{H}$. We then have

$$f(x) = \int d\mu(y)\, e^{i(x,y)}. \tag{3.32}$$

By $||\mu||$ me denote the supremum of the measure μ. Next we write $\mathcal{F}(\mathbb{R}^n)$ for the space of functions on $\mathcal{H}$ which are Fourier transforms of bounded complex measures. We define a norm for a function $f \in \mathcal{F}$ as $||f||_0 = ||\mu||$, where μ is

the measure corresponding to the function f. Note that this implies that f is a bounded continous function in $\mathcal{H}$.

The main definition is the following: for a function $f \in \mathcal{F}$,

$$\frac{1}{N}\int dx\, e^{\frac{i}{2}|x|^2} f(x) = \int d\mu(x)\, e^{-\frac{i}{2}|x|^2}. \tag{3.33}$$

With this notation, we then define the *Fresnel integral* as

$$\mathcal{F}(f) := \frac{1}{N}\int dx\, e^{\frac{i}{2}|x|^2} f(x). \tag{3.34}$$

From the above equations we get that $\mathcal{F}(f)$ is a continous bounded functional. Albeverio and Høegh-Krohn then prove that

$$\begin{aligned} \mathcal{F}(1) &= 1 \\ |\mathcal{F}(f)| &\leq ||f||_0 \end{aligned}$$

as well as that the Fresnel integral supports a functional integral version of the Fubini theorem and that the Fresnel integrals are invariant under transformations $x \to Ox + a$ of the Hilbert space where O is an orthogonal transformation from $\mathcal{H}$ to $\mathcal{H}$ and a is an arbitrary element of the Hilbert space.

The paper then goes on to define the same basic properties stated above for a general Hilbert space which is separable. The following definition is superficially almost identical with Ito's definition which we encountered in section 3.3.1. The connection between these two is of the form

$$\mathcal{F}(f) = \frac{1}{N}\int dx\, e^{\frac{i}{2}|x|^2} f(x) = F(e^{\frac{i}{2}|x|^2} f),$$

where F is a function defined as a limit (in fact by the equation (3.18)).

The general definition given above is then used on the real Hilbert space $\mathcal{H}$ of real continous functions $\gamma(\tau)$ from $[0,t]$ to $\mathbb{R}^n$ such that $d\gamma/d\tau \in L^2$. We also define $\gamma(t) = 0$ and the inner product on $\mathcal{H}$ as

$$(\gamma_1, \gamma_2) = \int_0^t d\tau\, \frac{d\gamma_1}{d\tau}\frac{d\gamma_2}{d\tau}. \tag{3.35}$$

(Notice the similarity with DeWitt-Morette's definition.)

This definition is then used to give the main result, a proposition called the *Feynman-Ito formula*, which is the quantum mechanical version of the Feynman-Kac formula of Wiener integrals.

Theorem 3.1 (The Feynman-Ito formula) *Let V and ϕ be Fourier transforms of bounded complex measures on $\mathbb{R}^n$. Let $\mathcal{H}$ be the real Hilbert space of continous paths $\gamma(\tau)$ from $[0,t]$ to $\mathbb{R}^n$ such that $\gamma(t)=0$ and $\dot{\gamma}\in L^{2,1}$ with inner product $(\gamma_1,\gamma_2)=\int_0^t d\tau\ \dot{\gamma}_1\dot{\gamma}_2$. Then*

$$f(\gamma)=e^{-i\int_0^t d\tau\ V(\gamma(\tau)+x)}\phi(\gamma(0)+x)$$

is in $\mathcal{F}(\mathcal{H})$, the space of Fresnel integrable functions on $\mathcal{H}$ and the solution of the Schrödinger equation

$$i\frac{\partial\psi}{\partial t}=-\frac{1}{2}\nabla^2\psi+V\psi$$

with the boundary condition $\psi(x,0)=\phi(x)$ is given by the normalized integral

$$\begin{aligned}\psi(x,t) &= \frac{1}{N}\int d\gamma\, e^{\frac{i}{2}|\gamma|^2} f(\gamma)\\ &= \int d\gamma\, e^{i\int_0^t d\tau\left[|\frac{d\gamma}{d\tau}|^2-V(\gamma(\tau))\right]}\phi(\gamma(0)+x).\end{aligned}\tag{3.36}$$

If the potential is of the quadratic form $V=\frac{1}{2}xAx$, where A is a strictly positive definite form (i.e. matrix) on $\mathbb{R}^n$, the above formula does not hold. The Lagrangian with such a potential term is not Fresnel integrable. However, the result can be generalized for such potentials. In this case we must apply the condition

$$|\gamma|^2=\int_0^t d\tau\ (\dot{\gamma}^2+\gamma^2)<\infty$$

and the inner product

$$<\gamma,\gamma>=\int d\tau\ (\dot{\gamma}^2-\gamma A^2\gamma)$$

on the paths to get the result that the same Feynman-Ito formula also holds in this case.

Although the above definition is formally consistent it does have its limitations. For example, the space of paths is assumed to have $\int d\tau|\dot{\gamma}|^2<\infty$, which – as in the case of pseudomeasures – is contrary to the intuitive view that the paths which are essentially nondifferentiable contribute to the path integral. So in principle this definition misses some of the generality of a proper definition, at least if we compare it to the Wiener integral. The interpretation of the term $\dot{\gamma}$ is a very difficult and is left a little ambiguous.

3.3.4 Polygonal paths

The third major contribution to the proper definition of Feynman path integrals is due to A. Truman. In his papers ([Tru76],[Tru77]) he introduces a definition based on the work of Albeverio and Høegh-Krohn as well as that of DeWitt-Morette. The concept of *polygonal paths* essentially only extends the space of functionals for which the integral (3.36) exists and thus making the Feynman-Ito formula a more useful tool. The papers greatly clarify the preceeding work of the above-mentioned authors.

To introduce this approach to path integrals we begin by assuming that X is a separable reflexive Banach space. This implies the existence of a dual space X' with the action of an element $x' \in X'$ defined as $< x', x >$. As with Fresnel integrals, we then define the integrable functions as Fourier transforms of bounded complex measures on X'. The integrable functions $f(x)$ are thus of the form

$$f(x) = \int d\mu(x')\, e^{-i<x',x>}.$$

Let $W(x',x')$ now be a variance such that $\exp\{-(i/2)W(x',x')\}$ is measurable. We then define the Feynman path integral as

$$I(f) = \int d\mu(x')\, e^{-\frac{i}{2}W(x',x')}. \tag{3.37}$$

This is almost an exact synthesis of the definitions by DeWitt-Morette and Albeverio and Høegh-Krohn. In non-relativistic quantum mechanics we can choose X to be a Hilbert space. On such space $W(x',x') = ||x'||^2$ and the equation (3.37) is identical with the equation (3.34).

Having defined the path integral we still are posed with the same problem: what is the interpretation of the derivative $d\gamma/d\tau$? Truman proposes two alternative ways to understand the role of such derivative: firstly, we may be able to isolate the functions for which $d\gamma/d\tau$ does not exist and then define the path integral on a quotient space. Secondly, we could interpret the derivative "loosely" as a weak derivative. Truman himself chooses the latter option.

Let us take H as the space of real-valued continous functions $\gamma(\tau) : [0,t] \to \mathbb{R}$ with $\frac{d\gamma}{d\tau} \in L^2$ and the end point of every path is fixed, $\gamma(t) = 0$. We begin by noticing that any $\gamma \in H$ can then be written in the form

$$\begin{aligned}\gamma(\tau) = {} & \alpha_0(\tau - t) - \sum^{\infty} \frac{\alpha_n t}{2\pi n} \sin\left(\frac{2\pi n\tau}{t}\right) \\ & + \sum^{\infty} \frac{\beta_n t}{2\pi n}\left[1 - \cos\left(\frac{2\pi n\tau}{t}\right)\right], \quad \tau \in [0,t].\end{aligned} \tag{3.38}$$

$\alpha_n, \beta_n \in \mathbb{R}$ are the usual Fourier coefficents of $\frac{d\gamma}{d\tau}$. We apply the condition $\sum^{\infty}(\alpha_n^2 + \beta_n^2) < \infty$ on the coefficents α_n, β_n. Note that in the above formula and

in the following we have $\tau \in [0,t]$. With this definition we instantly know that the derivative $\gamma'(\tau)$ exists and that it can be explicitely written down.

We can now write write down the inner product on the space H of paths $\gamma(\tau)$. It is

$$(\gamma,\gamma') := \int d\tau \, \frac{d\gamma}{d\tau}\frac{d\gamma'}{d\tau} = t\alpha_0\alpha_0' + \frac{t}{2}\sum^{\infty}\alpha_n\alpha_n' + \frac{t}{2}\sum^{\infty}\beta_n\beta_n'. \tag{3.39}$$

H is then a real separable Hilbert space with reproducing kernel $G(\sigma,\tau) = (t - \max(\sigma,\tau))$. The reproducing property is stated by the equation

$$\gamma(\sigma) = \big(G(\sigma,\tau),\gamma(\tau)\big), \quad \sigma \in [0,t]. \tag{3.40}$$

We recall that $G(\sigma,\tau)$ is the Green's function of the differential operator $-\frac{d^2}{d\tau^2}$ with boundary conditions $\frac{dG}{d\tau}(\sigma,\tau=0) = 0, G(\sigma,\tau=t) = 0$.

Next we introduce the linear map $P_n : H \to H$ by

$$(P_n\gamma)(\tau) = \sum_{j=0}^{n-1}\big[G(\frac{(j+1)t}{n},\tau) - G(\frac{jt}{n})\big][\gamma_{j+1} - \gamma_j]\frac{n}{t} \tag{3.41}$$

where $\gamma_j = \gamma(jt/n), j = 0,1,\ldots,n$. The reproducing property of the kernel $G(\sigma,\tau)$ implies

$$(\gamma', P_n\gamma) = \sum_{j=0}^{n-1}(\gamma'_{j+1} - \gamma'_j)(\gamma_{j+1} - \gamma_j)\frac{n}{t} = (P_n\gamma',\gamma). \tag{3.42}$$

Thus $P_n^* = P_n$.

If we substitute $(t - \max(\sigma,\tau))$ for $G(\sigma,\tau)$ in the equation (3.41) we get

$$(P_n\gamma)(\tau) = \gamma_j + (\tau - \frac{jt}{n})(\gamma_{j+1} - \gamma_j)\frac{n}{t}, \quad \frac{jt}{n} \le \tau \le \frac{(j+1)t}{n}. \tag{3.43}$$

From the form of equation (3.43) we immediately see that $(P_n\gamma)$ is just the piecewise linear polygonal approximation to γ. It can be proved that $P_n^2 = P_n$ and thus P_n is indeed a projection. This procedure of taking the piecewise linear approximation of the path gives the name *polygonal path formulation* of Feynman path integrals to the alternative definition of functional integration which we will present in the following.

We first note that the polygonal paths $(P_n\gamma)$ are dense in H in the sense that if $V = \{\gamma \in H| \ \ ||P_n\gamma - \gamma|| \to 0 \text{ as } n \to \infty\}$ then $V = H$. For a proof see Truman [Tru76].

To connect the above formalism to the usual definition of path integrals and to the Feynman-Ito formula we introduce the *Feynman maps* $\mathcal{F}^s$. The label s refers

to a complex variable with Im $s \leq 0$. For the normal Feynman path integral $s = 1$ and then we write $\mathcal{F}^1 \equiv \mathcal{F}$.

Let us define the complex Gaussian $e_s : H \to \mathbb{C}$ as

$$e_s[\gamma] = e^{\frac{i}{2s}(\gamma,\gamma)}. \tag{3.44}$$

With this notation we define the Feynman map $\mathcal{F}^s[f]$ of the complex-valued functional f. We first denote $\mathcal{F}_n^s[f]$

$$\mathcal{F}_n^s[f] = \int_{P_nH} d^n\gamma\,(fe_s)\circ P_n\Big[\int_{P_nH} d^n\gamma\,(e_s\circ P_n)\Big]^{-1}, \tag{3.45}$$

where integration is done from $-\infty$ to ∞ for each of the varibles $\Delta\gamma_j = (\gamma_{j+1} - \gamma_j)$. The equation (3.45) is normalized so that $\mathcal{F}_n^s[1] = 1$. The Feynman map $\mathcal{F}^s[f]$ is defined by the limit $\mathcal{F}^s[f] = \lim_{n\to\infty} \mathcal{F}_n^s[f]$ when it exists. We also write $f \in \mathcal{F}^s(P_\infty H)$ if and only if the above limit exists. $\mathcal{F}^s(P_\infty H)$ is the class of integrable functionals.

We now use the definition of Albeverio and Høegh-Krohn that $\mathcal{F}(H)$ is the space of functionals which are Fourier transforms of complex-valued measures on H which are of bounded absolute variation. Each $f \in \mathcal{F}(H)$ can be written in the form of the equation (3.32). Truman proves that $\mathcal{F}(H) \subset \mathcal{F}^s(P_\infty H)$ and that if $f \in \mathcal{F}(H)$

$$\mathcal{F}^s[f] = \int_H d\mu_f(\gamma)\, e^{-\frac{is}{2}(\gamma,\gamma)}. \tag{3.46}$$

Hence for Im $s \leq 0$ we have

$$|\mathcal{F}^s[f]| \leq \int_H |d\mu_f(\gamma)| = ||f||_0 < \infty. \tag{3.47}$$

If we compare this with the results of Albeverio and Høegh-Krohn shown in section 3.3.3 we see that $\mathcal{F}^s(H)$ is an extension of the original space of integrable functions.

If we take $s = -i$ we have the Wiener integral, since

$$e_i[\gamma] = e^{-\frac{1}{2}(\gamma,\gamma)}, \tag{3.48}$$

with (γ,γ) given by $||\dot{\gamma}||^2$. If we compare the functional $e_i[\gamma]$ of the above equation we find that it is the Wiener measure. Truman uses this connection to establish some important properties of the general theory: the translation formula for linear transformations $\gamma \to \gamma + a$ with $a \in H$.

What remains to do is to connect the polygonal path formalism to the usual definition of the path integral and to show the validity of the Feynman-Ito formula for path integrals defined by polygonal paths. We begin by using the notation $(P_n\gamma + X)$ for the polygonal path

$$(P_n\gamma + X)(\tau) = (P_n\gamma)(\tau) + X, \quad \tau \in]0,t[, \tag{3.49}$$

where X is constant. We let $\mathcal{S}[P_n\gamma + X]$ denote the classical action. Explicitely it is, as usual,

$$\mathcal{S}[P_n\gamma + X] = \sum^{n-1} \frac{(\gamma_{j+1} - \gamma_j)^2}{2\Delta t} - \int_0^t d\tau\, V[P_n\gamma + X], \tag{3.50}$$

with $\Delta t = t/n$.

The Feynman-Ito formula is then written in the normal way as the solution to the Schrödinger equation with initial data $\phi = \int \exp(i\alpha X) d\nu(\alpha)$ and with a real-valued potential $V = AX^2 + BX + C + \int \exp(i\alpha X) d\mu(\alpha)$. ν and μ are complex measures of bounded variation on $\mathbb{R}$. The solution is

$$\psi(X,t) = \mathcal{F}\,[e^{-i\int_0^t d\tau\, V[\gamma(\tau)+X]} \phi[\gamma(0) + X]]. \tag{3.51}$$

Truman also shows in [Tru77] that his formalism behaves nicely as far as semiclassical expansions are conserned. The limit $\hbar \to 0$ is also derived explicitely and shown to give the desired result – classical mechanics results when one extremizes the action in the integral (3.51).

Although the polygonal path formulation of the Feynman path integrals is clearly an extension of the aforementioned definitions of DeWitt-Morette and Albeverio and Høegh-Krohn, it too has its limitations. For example, even though the class of potentials that can be handled with this formalism is wider and the procedure is far more lucid than in the formalism of Albeverio and Høegh-Krohn when it comes to non-singular quadratic potentials, we still cannot apply this formalism to every possible case. For example, the general polynomial potential is not integrable according to the polygonal path formulation of path integration.

Another point that must be criticized is the formalism itself; when we use the polygonal path approximation we are in fact using a discretized version of the path integral and using the limit $n \to \infty$ as the definition. This is very evident in equation (3.50). Although the limit exists we are still short of a definition on the path of spaces itself, and not merely on a limited subset of such a space.

3.4 Discretization of Feynman path integrals

Even though it is not mathematically correct it is quite possible to take the original definition of the Feynman path integral and use it as if it was a well-founded

construct. One then defines the functional "integral"

$$I = \int \mathcal{D}x\, e^{i\mathcal{S}}$$

as the *formal* limit $N \to \infty$ of the finite-dimensional integral

$$I_N = \int \cdots \int \prod_{i=1}^{N-1} dx_i \prod_{i=1}^{N} \Big(\frac{1}{\sqrt{2\pi i\varepsilon}} \Big) e^{i\sum_{i=1}^{N} \varepsilon \mathcal{L}(x_i, x_i/\varepsilon)}, \tag{3.52}$$

where $\mathcal{L}$ is the Lagrangian of classical mechanics. This is an appealing and widely-used approach to path integrals since the explicit discretization is very easy to implement to calculation of such integrals on a computer.

One must however be careful when one uses this definition, since it explicitely depends on the way one writes the discretized integral I_N. We have already seen in the case of Wiener integrals how sensitive the integral can be to the choice of the discretization rule. In fact, one should always provide a discretization rule to the equation (3.52) when one uses this approach. If such rule is not given, the integral does not have a unique meaning – nor value. Note that discretization does not directly have anything to do with quantization.

Fortunately a lot of research has been done on this approach to path integration and on how discretization rules affect the path integral. Ever since the appearance of Feynman path integrals physicist have developed methods of calculating these integrals via discretization. It is still the most important way of actually computing the value of a path integral. Next we will discuss the discretization procedure a bit more formally, following closely Langouche et al [LRT82].

Let us now define more clearly what we mean by discretization. We approach the subject with phase space integration, since it is more general and in most cases can easily be reduced to a configuration space integral by integration over the momentum varibles. Let us consider the phase space functional integral

$$I = \int_{\gamma} \mathcal{D}q\mathcal{D}p\, \exp\Big(i \int_{t_0}^{t} d\tau [p \cdot \dot{q} - H(p(\tau), q(\tau), \tau)] \Big)\, \{\delta\}, \tag{3.53}$$

where γ stands for certain discretization rule such that for a given function h^{γ}

$$h^{\gamma}(p, q, q', \tau) \approx H(p, q, \tau) \tag{3.54}$$

in the sense that the discretized integral

$$I_N = \int \Big(\prod_{i=0}^{N} dq_i \prod_{j=1}^{N} \frac{dp_j}{2\pi i} \Big) \exp\Big(i \sum_{j=1}^{N} \varepsilon [p_j \cdot \frac{\Delta q_j}{\varepsilon} - h^{\gamma}(p_j, q_j, q_{j-1}, \bar{t}_j)] \Big)\, \{\delta\}$$

has the limit $I = \lim_{N\to\infty} I_N$. $\{\delta\}$ is a formal notation which represents the constraints on the coordinates – it is of the form $\delta(q_0 - 0)\delta(q_t - x)$ and makes sure that the starting and end points of the path are fixed. $\bar{t}_j \in [t_{j-1}, t_j]$ and $\Delta q_j = q_j - q_{j-1}$. We also generally write $\dot{q}_j := \Delta q_j/\varepsilon$.

Essentially the function $h^\gamma(p, q', q, \tau)$ thus defines the discretization γ. We also note that even though we have not explicitely written $\hbar$ in the above equations the function h^γ can depend on it.

It is possible that two discrete approximations to (3.53) yield the same value for the integral. We say then that two discretizations γ' and γ'' are equivalent if the limit $I = \lim_{N\to\infty} I_N$ is same for both. Note that this equivalence may be restricted to a certain set of functions $H(p, q, \tau)$. We can write explicitely the conditions of equivalence for two discretizations. Let us consider a classical Hamiltonian function which is quadratic in p. The discretized Hamiltonian is then of the form

$$h^\gamma(p, q', q, \tau) = \frac{1}{2} p_\mu p_\nu g^{\mu\nu}_{[\gamma]}(q', q, \tau) - p_\mu a^\mu_{[\gamma]}(q', q, \tau) + v_{[\gamma]}(q', q, \tau), \tag{3.55}$$

where $g^{\mu\nu}$ is the metric of the underlying space.

Let us denote $\Delta := q' - q$. Then the conditions

$$\begin{aligned} g^{\mu\nu}_{[\gamma]} - g^{\mu\nu}_{[\gamma']} &= O(\varepsilon^{3/2}) = O(\Delta^3) \\ a^\mu_{[\gamma]} - a^\mu_{[\gamma']} &= O(\varepsilon) = O(\Delta^2) \\ v_{[\gamma]} - v_{[\gamma']} &= O(\varepsilon^{1/2}) = O(\Delta) \end{aligned} \tag{3.56}$$

are sufficent for the discretizations γ and γ' to be equivalent. For a proof and further details see Langouche et al [LRT82].

Let us now introduce a general notation for the most commonly used discretization rules. They depend on one parameter $\alpha \in [0, 1]$. If

$$q^{(\alpha)} = q + \alpha(q' - q) \tag{3.57}$$

then $\gamma_1(\alpha)$ is defined by

$$h^{\gamma_1(\alpha)}(p, q', q, \tau) := H(p, q^{(\alpha)}, \tau). \tag{3.58}$$

We define $\gamma_2(\alpha)$ by

$$h^{\gamma_2(\alpha)}(p, q', q, \tau) := (1 - \alpha)H(p, q, \tau) + \alpha H(p, q', \tau). \tag{3.59}$$

One can easily verify that these both satisfy the property (3.54) for any α. If we explicitely write down the discretizations $\gamma_1(0)$, $\gamma_1(\frac{1}{2})$ and $\gamma_1(1)$ we find

that they are the prepoint-, midpoint- and postpoint-prescriptions encountered in analysis of stochastic integration discussed in section 2.3.

Note that a linear combination of the above discretization rules is also a discretization rule.

Although the above definition used Hamiltonian functions and the phase space formulation of functional integrals the same notation and definitions apply also to configuration space path integrals. We will then have to define the discretization on several functions rather than only the Hamiltonian; for the discretized Lagrangian function

$$\frac{1}{2}g_{\mu\nu}^{[\gamma]}(q_j,q_{j-1},\bar{t}_j)\frac{\Delta q_j^\mu}{\varepsilon}\frac{\Delta q_j^\nu}{\varepsilon}+a_\mu^{[\gamma]}(q_j,q_{j-1},\bar{t}_j)\frac{\Delta q_j^\mu}{\varepsilon}-v^{[\gamma]}(q_j,q_{j-1},\bar{t}_j)$$

we must give the functions $g_{\mu\nu}^{[\gamma]}, g^{[\gamma]}, a_\mu^{[\gamma]}$ and $v^{[\gamma]}$ in order to define the discretization. The same conditions (3.56) ensure the equivalence of two discretizations. We can use the $\gamma_i(\alpha)$-notation on these functions by the generalization

$$f^{\gamma_1(\alpha)}(q,q',\tau)=F(q^{(\alpha)},\tau)$$

and analogously on $\gamma_2(\alpha)$. An example of use of these discretization rules on configuration space integrals is the calculation of a Wiener integral in section 2.4.2, which also shows that the notation is also useful in Wiener integrals.

The discretization rules are generated by the use of operator formalism to define path integrals. When we consider the short-time propagator we can shift $\widehat{p}$- and $\widehat{q}$-operators with respect to each other according to the basic commutation relation

$$[\widehat{q},\widehat{p}]=i. \tag{3.60}$$

To see this better let us consider the Hamiltonian operator

$$\widehat{H}(\widehat{p},\widehat{q})=-\frac{1}{2}\widehat{p}^2G(\widehat{q})-\widehat{p}A(\widehat{q})+V(\widehat{q}). \tag{3.61}$$

This operator can be written, for any value $\alpha\in[0,1]$, in the form

$$\begin{aligned}\widehat{H}(\widehat{p},\widehat{q}) &= (1-\alpha)[-\frac{1}{2}\widehat{p}^2G(\widehat{q})-\widehat{p}A(\widehat{q})]\\ &+ \alpha[-\frac{1}{2}G(\widehat{q})\widehat{p}^2-A(\widehat{q})\widehat{p}]\\ &+ \alpha[-\frac{1}{2}[\widehat{p}^2,G(\widehat{q})]-[\widehat{p},A(\widehat{q})]+V(\widehat{q}).\end{aligned} \tag{3.62}$$

By using the commutation relation (3.60) we get for the arbitrary operator $F(\widehat{q})$

$$\begin{aligned}\left[\widehat{p},F(\widehat{q})\right] &= -iF'(\widehat{q})\\ \left[\widehat{p},F(\widehat{q})\right] &= -2i[(1-\alpha)\widehat{p}F'(\widehat{q})+\alpha F'(\widehat{q})\widehat{p}]-2(\alpha-\frac{1}{2})F''(\widehat{q}),\end{aligned} \quad (3.63)$$

where $F' := \partial F/\partial q$. Applying this result to the operator $\widehat{H}(\widehat{p},\widehat{q})$ we get, by using the results of section 3.2.2, that the path integral corresponding to the propagator $\langle q|\exp(i\widehat{H})|q_0\rangle$ can be written in the form

$$I = \int_{\gamma_2(\alpha)} \mathcal{D}q\mathcal{D}p\ \exp\left(i\int_{t_0}^{t} d\tau[p\dot{q}-H^{\alpha}(p,q)]\right)\ \{\delta\}, \quad (3.64)$$

with

$$\begin{aligned}H^{\alpha} &= -\frac{1}{2}p^2G(q)-p[A(q)+\alpha G'(q)]\\ &\quad + V(q)-\alpha A'(q)+\alpha(\alpha-\frac{1}{2})G''(q).\end{aligned} \quad (3.65)$$

Thus, by using commutation relations on the operator (3.61) we have explicitely introduced terms in the path integral which depend on the form of the operator. Therefore also the value of the path integral depends on the definition of the integral. Definition in this case means that we must define what we mean by the operator $\widehat{(}\widehat{p},\widehat{q})$. Since the propagator $\langle q|\exp(i\widehat{H})|q_0\rangle$ must be independent of any discretization we deduce that the α-dependence in the integrand cancels out any α-dependence of the functional integral.

Looking at the equation (3.65) we see that the potential term $V(q)$ does not depend on α. This is due to the fact that

$$\langle q_j|V(\widehat{q})|q_{j-1}\rangle = V(q_j)\langle q_j|q_{j-1}\rangle = V(q_{j-1})\langle q_j|q_{j-1}\rangle.$$

One can easily see that $\gamma_1(\alpha)\equiv\gamma_2(\alpha)$ when $\alpha=0$ or $\alpha=1$. The discretizations are also equal whenever H is a function of the form

$$H = f(q)p^2+g(q),$$

where $f(q)$ is a linear function $f(q)=aq+b$ of q with $a,b\in\mathbb{R}$.

We also make the observation that $\overline{t}_j\in[t_j,t_{j-1}]$ can be chosen arbitrarily. This can be seen by expanding the path integral according to two different discretization rules and comparing the terms corresponding to the potential term.

3.4.1 Connection to the operator ordering problem

If we use the definition of last section another problem will occur; when we discretize the functional integral and write the short-time propagator like in section 3.2.2, we have to decide what is the interpretation of the general Hamiltonian operator

$$\widehat{H} = H(\widehat{q}, \widehat{p}, t) \tag{3.66}$$

in terms of the discretized varibles q_i and p_i. (H is the classical Hamiltonian function.) The path integral does not know anything about operators and hence about their ordering. Everything is classical in the path integral.

Operator ordering is not trivial since, for example, there are many different forms of Hamiltonian operator which correspond to the classical quantity p^2q^2. The noncommutativity of $\widehat{q}$ and $\widehat{p}$ demands that we describe a definite order in which the operators corresponding to the classical quantities are written. The above Hamiltonian function can be written as $\widehat{q}\widehat{p}^2\widehat{q}$, $\widehat{q}^2\widehat{p}^2$ or $\widehat{p}^2\widehat{q}^2$, for example. One can also always use a linear combination of the above operators.

This problem is called the *operator ordering problem*. It is quite common in quantum mechanics. And since the path integral formalism is manifestly based on classical quantities that represent quantities in quantum mechanics, the problem is also present in path integration when we calculate the probability amplitude via the short-time operator formalism. We must then make the decision of which ordering rule we use when we make the quantization $H(q,p) \to \widehat{H}(\widehat{q}, \widehat{p})$.

The operator ordering problem in connection to path integration has been studied, for example, by Cohen [Coh66],[Coh70] and Mehta [Meh64]. The analysis is based on the fact that we a function $\Omega(u,v)$ which maps the classical Hamiltonian function $H(q,p)$ into a Hamiltonian operator $\widehat{H}(\widehat{q}, \widehat{p})$. Let us first define what we mean by a corresponce rule:

Definition 3.1 (Correspondence rule) *A correspondence rule is a linear mapping from phase space functions $B(q,p)$ into operators $\widehat{B}(\widehat{q}, \widehat{p})$. It is completely characterized by the way it acts on the phase space function* $\exp(iu \cdot q + iv \cdot p)$, *where $u \cdot q$ is the inner product of the space in question. In $\mathbb{R}^n$, for example, we have $u \cdot q := \sum^n u_i q_i$, $v \cdot p := \sum^n v_i p_i$.*

All usual correspondence rules are of the form

$$\exp(iu \cdot q + iv \cdot p) \to \Omega(u,v) \exp(iu \cdot \widehat{q} + iv \cdot \widehat{p}). \tag{3.67}$$

In the following we shall take $n = 1$ and so our phase space is $\mathbb{R}^2$.

$\Omega(u,v)$ is an analytical function of the components of u and v. A list of the usual correspondence rules with their Ω-function is given in table 3.1, which also

shows the form of the operator the function $q^n p^m$ is mapped to. The names of the most commonly used correspondence rules are also shown in the table. These names originate from the operator formulation of quantum mechanics.

Correspondence rule	$\Omega(u,v)$	Mapping of $q^n p^m$
Weyl	1	$2^{-n}\sum_{l=0}^{n}\binom{n}{l}\widehat{q}^{n-l}\widehat{p}^{m}\widehat{q}^{l}$
Symmetric	$\cos(\frac{1}{2}u\cdot v)$	$\frac{1}{2}(\widehat{q}^n\widehat{p}^m+\widehat{p}^m\widehat{q}^n)$
Standard	$\exp(-\frac{1}{2}iu\cdot v)$	$\widehat{q}^n\widehat{p}^m$
Anti-standard	$\exp(+\frac{1}{2}iu\cdot v)$	$\widehat{p}^m\widehat{q}^n$
Born-Jordan	$\sin(\frac{1}{2}u\cdot v)/(\frac{1}{2}u\cdot v)$	$\frac{1}{m+1}\sum_{l=0}^{n}\widehat{p}^{m-l}\widehat{q}^{n}\widehat{p}^{l}$

Table 3.1: List of operator ordering rules

Note that if we consider the mapping $H\rightarrow\widehat{H}$ by the correspondence rule, the simple cases $f(q)\rightarrow f(\widehat{q})$ and $g(p)\rightarrow g(\widehat{p})$ also lead to the restrictions $\Omega(u,0)=1$ and $\Omega(0,v)=1$, respectively. Thus, if H is of the form $f(q)+g(p)$ then all the correspondence rules give the same operator.

We also define the inverse mapping from the operator $\widehat{H}$, ordered by some ordering rule, to the ordinary function H as

$$e^{i(u\widehat{q}+v\widehat{p})}\rightarrow\Omega^{-1}(u,v)e^{i(uq+vp)}. \tag{3.68}$$

This is of great interest to us since if we use the short-time propagator approach we will have to determine the function corresponding to the Hamiltonian operator in the discretized form. We will denote this inverse mapping with θ_Ω and drop the subscript whenever it can be done without confusion.

Let us now give the explicit formulae for the mappings $B\rightarrow\widehat{B}^{\Omega}$ and $\widehat{B}\rightarrow B^{\theta}$. Let $\tilde{B}(p,q)$ be the two-dimensional Fourier transform of the function $B(u,v)$:

$$\tilde{B}(p,q)=\mathcal{F}\{B(p,q)\}=\frac{1}{(2\pi)^2}\int du\,dv\,B(p,q)e^{-i(uq+vp)}. \tag{3.69}$$

Then the operator $\widehat{B}(\widehat{p},\widehat{q})$ corresponding to a specific correspondence rule $\Omega(u,v)$ is given by

$$\widehat{B}(\widehat{p},\widehat{q})=\int du\,dv\,\tilde{B}(u,v)\Omega(u,v)e^{i(u\widehat{q}+v\widehat{p})}. \tag{3.70}$$

This can be referred to as the *operator Fourier transform*. With the help of equation (3.69) we can write the above equation in the form

$$\widehat{B}(\widehat{p},\widehat{q})=\int du\,dv\,dp\,dq\,B(q,p)\Omega(u,v)e^{-iu(q-\widehat{q})-iv(p-\widehat{p})}. \tag{3.71}$$

To construct the inverse mapping from $\widehat{B}(\widehat{p},\widehat{q})$ to $B(p,q)$ we begin by defining $\widehat{B}(\widehat{p},\widehat{q})$ as

$$\widehat{B}(\widehat{p},\widehat{q}) = \int du\, dv\, G(u,v) e^{i(u\widehat{q}+v\widehat{p})}, \tag{3.72}$$

where $G(u,v)$ is the Fourier transform

$$G(u,v) = \frac{1}{2\pi}\mathrm{Tr}[\widehat{B}(\widehat{p},\widehat{q}) e^{-i(u\widehat{q}+v\widehat{p})}].$$

(Tr denotes the trace $\int dq\, \langle q|\widehat{B}(\widehat{p},\widehat{q})|q\rangle$.)

We denote the inverse $\Omega^{-1}(u,v)$ of Ω as θ_Ω. Whenever we can we drop the subscript and simply use the notation θ. We then define

$$B^\theta(p,q) = \int du\, dv\, G(u,v)\Omega^{-1}(u,v) e^{i(uq+vp)}. \tag{3.73}$$

The quantity $B^\theta(p,q)$ is called the θ-*symbol* of the operator $\widehat{B}(\widehat{p},\widehat{q})$. (This terminology is due to Berezin [Ber80].)

We can also write $B^\theta(p,q)$ explicitely in terms of the operator $\widehat{B}(\widehat{p},\widehat{q})$ and the correspondence rule $\Omega(u,v)$:

$$B^\theta(p,q) = \int du\, dv\, \mathrm{Tr}[\widehat{B}(\widehat{p},\widehat{q})]\Omega^{-1}(u,v) e^{iu(q-\widehat{q})+iv(p-\widehat{p})}. \tag{3.74}$$

Two different mappings B^{θ_1}, B^{θ_2} of the operator $\widehat{B}$ are connected by the equation

$$B^{\theta_2}(p,q) = \Omega_1(-i\frac{\partial}{\partial q}, -i\frac{\partial}{\partial p})\Omega_2^{-1}(-i\frac{\partial}{\partial q}, -i\frac{\partial}{\partial p}) B^{\theta_1}(p,q) \tag{3.75}$$

What we need to do now is to show that the connection between the above correspondence functions and the discretization rules $\gamma_i(\alpha)$ introduced before. We begin by considering the Weyl correspondence rule (with $\Omega(u,v) \equiv 1$). We can derive the following formula for the general operator $\widehat{B}(\widehat{p},\widehat{q})$:

$$\langle q'|\widehat{B}(\widehat{p},\widehat{q})|q\rangle = \frac{1}{2\pi}\int dp\, \exp[ip\Delta]\Omega(-i\frac{\partial}{\partial\overline{q}}, -\Delta) B^{\theta_\Omega}(p,\overline{q}) \tag{3.76}$$

with $\Delta := q'-q$ and $\overline{q} := \frac{1}{2}(q'-q)$. We define

$$b^\theta(p,q',q) := \Omega(-i\frac{\partial}{\partial\overline{q}}, -\Delta) B^{\theta_\Omega}(p,\overline{q}). \tag{3.77}$$

Note that if $\Omega(u,0) \equiv 1$ then

$$b^{\theta}(p,q',q) = B^{\theta_\Omega}(p,\overline{q}). \tag{3.78}$$

Let us consider the short-time propagator of a system with Hamiltonian operator $\widehat{H}(\widehat{p},\widehat{q})$. The propagator can be written in the approximate form

$$\langle q_j|U(t_j,t_{j-1})|q_{j-1}\rangle = \int \frac{dp_j}{2\pi} e^{i[p_j\Delta q_j - \varepsilon h^{\gamma}(p,q_j,q_{j-1},\overline{t}_j)]} \tag{3.79}$$

according to the discretization rules of section 3.4.

On the other hand, there is a θ-symbol $H^{\theta}(p,q,\tau)$ corresponding to the Hamiltonian operator $\widehat{H}$ through the correspondence rule $\Omega(u,v)$. As show in equation (3.77), there exists a discretized version of such symbol and that the propagator can alternatively be written as

$$\langle q_j|U(t_j,t_{j-1})|q_{j-1}\rangle = \int \frac{dp_j}{2\pi} e^{i[p_j\Delta q_j - \varepsilon h^{\theta}(p,q_j,q_{j-1},\overline{t}_j)]}. \tag{3.80}$$

Notice that equations (3.79) and (3.80) are almost identical; in fact, as we take the continuum limit $N \to \infty$ we write the limit as a functional integral such as the equation (3.53) with $H(p,q,\tau) = H^{\gamma(\theta)}(p,q,\tau)$, so that

$$H^{\gamma(\theta)}(p,q,\tau) = h^{\theta}(p,q',q,\tau), \tag{3.81}$$

in the same sense as the equation (3.54).

Notice, however, that $H^{\gamma(\theta)}$ is not necessarily equal to the θ-symbol H^{θ} of the operator $\widehat{H}$. If $\Omega(u,0) \equiv 1$ then this is true. Equation (3.81) shows that there is a connection between the discretization of the functional integral and the ordering rule imposed on the Hamiltonian operator. To be more specific, the correspondence rule $\Omega(u,v)$ and its inverse, the θ-symbol, induce a discretization h^{γ}, which depends explicitely on the chosen correspondence rule.

To make this connection clearer we introduce the mapping

$$\Omega(u,v,\alpha) = \exp[i(\frac{1}{2}-\alpha)uv], \tag{3.82}$$

which is related to the $\gamma_1(\alpha)$-discretization. For $\alpha = 0, \frac{1}{2}$ and 1 this coincides with anti-standard, Weyl and standard correspondence, respectively.

From equation (3.77) it follows that

$$\begin{aligned} h^{\theta_1(\alpha)}(p,q',q,\tau) &= \exp[-(\frac{1}{2}-\alpha)\Delta\frac{\partial}{\partial\overline{q}}]H^{\theta_1(\alpha)}(p,\overline{q}) \\ &= H^{\theta_1(\alpha)}(p,\overline{q}-(\frac{1}{2}-\alpha)\Delta) \\ &= H^{\theta_1(\alpha)}(p,q^{(\alpha)}). \end{aligned} \tag{3.83}$$

Correspondence rule	Discretization rule	Discretization prescription
Weyl	$\gamma_1(\frac{1}{2})$	mid-point
Symmetric	$\gamma_2(\frac{1}{2})$	
Standard	$\gamma_1(1) \equiv \gamma_2(1)$	post-point
Anti-standard	$\gamma_1(0) \equiv \gamma_2(0)$	pre-point
Born-Jordan	$\gamma(P)$	polygonal path

Table 3.2: Table of corresponding discretization and operator ordering rules

Since $\Omega_1(p,0,\tau) \equiv 1$ we have $H^{\gamma_1(\alpha)} \equiv H^{\theta_1(\alpha)}$. Other important correspondence rules and their associated discretizations are listed in the table 3.2.

Note that the Born-Jordan correspondence rule of table 3.2 refers to to the polygonal path formulation of path integrals discussed in section 3.3.4.

To end the discussion concerning the discretization of Feynman path integrals we recall that the discretization rule $\gamma_i(\alpha)$ affects the Feynman rules of quantum electrodynamics [LRT82]. Langouche et al also show that the discretization $\gamma_1(0)$, corresponding to the prepoint discretization prescription, yields the simplest Feynman rules.

The connection between operator ordering and discretization in functional integration has also been shown in path integrals defined without the limiting procedure. This has been done by Mizrahi [Miz78], [Miz79].

3.5 Feynman path integral on spaces with curvature

So far we have only dealt with path integrals which are defined on flat i.e. eucledian spaces corresponding to non-relativistic quantum mechanics. As such our current view is still quite limited since we should also be able to deal with cases where the underlying configuration space (or phase space) is a general manifold and possibly has some topological constraints – the movement of a particle may be restricted to some subspace by infinite potentials, for example. Another case where we need a more general concept of path integration is when we are discussing systems with rotational symmetry such as the Coulomb potential. The ultimate goal is to make the path integration method compatible with general relativity.

As one might expect, path integration is a much more complicated matter when we generalize it to non-eucledian spaces. Even the operator approach to quantum mechanics runs into trouble on spaces with curvature. Let us begin by reviewing the procedures necessary when we write down the path integral in a general coordinate system. This was first done by in B. S. DeWitt's classical paper [DeW57].

First we note that, in absence of a potential, the Schrödinger equation on the general manifold M is written as

$$i\frac{\partial\psi(q,t)}{\partial t} = -\frac{1}{2}\Delta_{LB}\psi(q,t). \tag{3.84}$$

Δ_{LB} is the *Laplace-Beltrami* operator

$$\Delta_{LB} := \frac{1}{\sqrt{g}}\partial_\mu(g^{\mu\nu}\sqrt{g}\partial_\nu),$$

where $g^{\mu\nu}$ is the metric tensor and $g = \sqrt{\det g^{\mu\nu}}$. Since both $g^{\mu\nu}$ and g depend on the coordinates q_μ we must be careful with operator ordering. Even though equation (3.84) does not make a reference to operators the question is more transparent in the treatment of the problem via path integral methods.

The classical Lagrangian and Hamiltonian functions are

$$\begin{aligned} L &= \frac{1}{2}g_{\mu\nu}\dot{q}^\mu\dot{q}^\nu + A_\mu\dot{q}^\mu \\ H &= \frac{1}{2}g^{\mu\nu}(p_\mu - A_\mu)(p_\nu - A_\nu), \end{aligned} \tag{3.85}$$

with $p_\mu = g_{\mu\nu}\dot{q}^\nu + A_\mu$. A is a vector potential.

The normal correspondence rule we apply in the case of a general manifold is

$$\widehat{q}^\mu = q^\mu, \quad \widehat{p}_\mu = -i(\partial_\mu + \frac{1}{4}\partial_\mu \ln g).$$

The Hamiltonian operator corresponding to the Schrödinger equation (3.84) is written as

$$\widehat{H}(\widehat{p},\widehat{q}) = \frac{1}{2}g^{-\frac{1}{4}}\widehat{p}_\mu g^{\mu\nu}g^{\frac{1}{2}}\widehat{p}_\nu g^{-\frac{1}{4}} = \frac{1}{2}\widehat{p}_\mu g^{\mu\nu}\widehat{p}_\nu + \widehat{Q}, \tag{3.86}$$

where

$$\widehat{Q} = -\frac{1}{2}g^{\frac{1}{4}}\Delta_{LB}g^{-\frac{1}{4}}.$$

Note that we have ordered the coordinates and momenta so that the covariance of the quantum theory is ensured.

What we need to do next is to find a propagator – in path integral form – which satisfies the Schrödinger equation (3.84). An often used guess is to use short-time propagators and seek for the correct limit via the limit $\hbar \to 0$ – the semiclassical approximation. This should yield

$$K = D^{1/2}e^{i\mathcal{S}}, \tag{3.87}$$

where D is a quantity which depends on the second derivatives of the classical action $\mathcal{S}$.

The main result of DeWitt is that the above guess does not give correct results. Instead, the propagator is of the form

$$\langle q'',t''|q',t'\rangle = \lim_{N\to\infty}(2\pi i\varepsilon)^{-1/2}\int\prod_{j=1}^{N-1}\Big[\frac{\sqrt{g(q_j)}dq_j}{(2\pi i\varepsilon)^{n/2}}\Big] \times \exp\Big(i\sum_{j=1}^{N}[S(q_j,t_j)+\frac{1}{12}R(q_j)]\Big), \tag{3.88}$$

where $\varepsilon = t_j - t_{j-1}$ and $S(q_j,t_j)$ is the classical action.

The most interesting term in the equation (3.88) is $\frac{1}{12}R$. R is the curvature scalar of the manifold. This term shows up as we do the coordinate transformations to the integration measure and action. Dimension analysis tells us that it has the dimension of energy, therefore it is possible to speak of the effective Hamiltonian operator

$$\widehat{H}_{eff} = \widehat{H} + \frac{1}{12}R(\widehat{q}).$$

This is an interesting and much disputed result. Several authors have derived the same result – except that the constant multiplying the curvature scalar is not $\frac{1}{12}$ as above. Since the curvature scalar directly affects the energy levels of the system it is alarming to note that there is a general disagreement on the value and the sign of the constant. However, we must recall that the magnitude of the curvature term is extremely small compared to the Hamiltonian. But, in order to make the theory of path integration compatible with general relativity we must know the exact value of the curvature term.

Kleinert has worked on the subject extensively; he has generalized the theory to spaces with curvature *and* torsion. The original formulation of DeWitt did not include this possibility. Kleinert has also applied the results of his work to physical systems, such as the hydrogen atom discussed in the next section. The results he has derived using path integrals are equal to the results obtained from the Schrödinger equation. A good account of the methods applicable to path integrals in spaces with curvatures can be found in Kleinert's book [Kle90].

Kleinert goes through the process of writing the propagator in terms of a general coordinate system. He argues that one must use a postpoint prescription if one wishes to obtain the correct result. Other important aspects of the procedure is to note that in the measure one should treat the differentials dq_i as differences Δq_i. This is reasonable since we are using the discretized version of the propagator. As with the Wiener integral (2.31), we must discretize the path integral in

eucledian space and after that do the coordinate transformations. With these assumptions, Kleinert finds that the multiplier of the curvature scalar is $-\frac{1}{6}$ instead of $\frac{1}{12}$. In order to prove that the result is correct he finds the solution to the path integral on the surface of a sphere by applying these methods and gets exactly the same result as obtained from the Schrödinger equation.

Other important observations include taking into account correct order of terms in the action. As we have seen, all terms which are of the order r^2/t contribute substancially to the result. And, as we do the coordinate transformations, the number of different terms usually rises and they become more complex. It may therefore sometimes be difficult to deduce the order of each term directly.

In order to be able to solve path integrals for systems with divergent potentials some authors have introduced the concept of *time transformations* to complement the methods of coordinate transformations in path integration. For example, the system with Coulomb potential is not directly solvable by normal path integral methods. Duru and Kleinert [D&K79] introduced *pseudotime* slicing, which actually is local reparametrization of time, to solve the Coulomb potential. Their method is now a basic tool when we solve systems with difficult potentials. Another reference which discusses time transformation methods of path integration – in more general terms – is by Young and DeWitt-Morette [Y&D86]. We note that this problem of *path collapse* would not appear if we were able to write the path integral as a true functional integral in continuum.

We also recall that the path integral formalism can also be easily applied to configuration spaces such as $SO(3)$. Problems involving the spin of particles often use such spaces.

3.5.1 The hydrogen atom

One particular example which shows that in some cases the path integral formulation of quantum mechanics is inferior to the traditional operator-based formulation is the hydrogen atom. The Feynman path integral formalism simply breaks down in the $1/r$ potential. Until very recently this simple system was unsolvable by methods of path integration, while the usual approach of angular momentum operators yields the exact eigenvalues and eigenfunctions of the hydrogen atom Schrödinger equation easily. Even Feynman himself could not solve this problem.

Finally, in 1979, Duru and Kleinert [D&K79] found the key to the solution of the path integral for hydrogen atom. The key was the aforementioned pseudotime slicing, which allowed them to work with the Coulomb potential. Another important factor in the solution was the use of Kustaanheimo-Stiefel transformation [K&S65], which transforms the problem of a Coulomb potential in three dimensions into a problem of harmonic oscillators in four dimensions. This al-

lows easier application of path integral methods.

In order to find correct results, one must also use careful analysis of how the path integral behaves on spaces with curvature.

Chapter 4

Conclusions

As we have seen, even after decades of mathematical and physical research a thoroughly rigorous mathematical theory for Feynman path integrals does not exist. Several good attempts have been made, but none of these is truly general enough to be *the* definition for path integration. All the alternative definitions discussed in this work seem to have to drawbacks: firstly, they are not intuitively so clear as Feynman's original definition. Secondly, they are not completely mathematically justified. This is quite clear if we compare them to the well-defined theory of Wiener integrals discussed earlier in this work. Even the style of this work shows this difference; while we were able to define the Wiener integral exactly and compactly with the language of mathematics we were compelled to use a more heuristic style in discussing the Feynman path integral.

The alternative definitions have turned out to be otherwise fruitful; mathematical methods based on them have usually found use in some particular application. However, as reviewed in this work, most of the attempts to properly define the Feynman path integral are rather abstract and it is easy to understand why they do not appeal to physicists whose work is mainly connected to the real physical world. The indirect techniques of defining the path integral by Fourier transforms are also quite non-intuitive.

Even the much-studied Brownian motion has its own ambiguities. It can be argued that stochastic integrals are not unique until we define the discretization rule imposed on the integral. Stratonovich and Ito integrals are both useful, and the choice of which to use depends on the problem. It should be noted that the theory does not prefer either. The value of a Wiener integral itself also depends on the choice of the discretization rule.

Much of the material cited in this work is quite old; most the of papers which discuss the definition of path integrals are from the end of the 1970's and the beginning of the 1980's. The era was a time of rapid development in functional integration and naturally some of the research was done on the the foundations of

the theory. However, only a very small number of articles have been written on the foundations of functional integration after the 1980's, at least if we compare it to the attention Feynman path integrals have otherwise recieved. This is an important point. While it is true that the Feynman path integrals were analyzed quite thoroughly and their limitations were brought to common knowledge, it seems that the majority of physicists and mathematicians have been at ease and concluded that the theory is perfect and concentrated their work on the calculational details and applications rather than the basic concepts.

Although the functional integration techniques seem to work well in almost all cases confronted by physicists today, the questions raised by the underlying mathematical problems still remain unsolved. However, the mathematical and physical research aimed at providing answers to these questions has in the process produced other important results. The connection between the operator ordering problem and the discretization rule used in path integration has been revealed, as well as the similarities between Browian motion, quantum mechanics and stochastic differential equations. One could say that functional integration as a mathematical field started to take its shape and unify as a result of the problems of Feynman path integrals.

Functional integration techniques have greatly improved due to the search for a rigorous theory. It is almost amazing how well functional integration works as a computational tool and as the mathematical language for quantum field theory. However, physicists should be more aware of the limitations of the path integral approach due to the lack of mathematical rigor.

In purely mathematical terms, the current basis of functional integration is intellectually and conceptually unsatisfactory. A proper and truly unifying mathematical foundation of functional integration should be found. A lot of mathematical research still remains to be done. As a side effect of this study we should expect to find more interesting connections and applications for functional integration and and simultaneously develop even more refined calculational methods of functional integration.

Bibliography

[A&H76] Albeverio, S., Høegh-Krohn, R.: Mathematical Theory of Feynman Path Integrals. Springer-Verlag 1976

[A&W81] Abbott, L. F., Wise, M. B.: Am. J. Phys. **49**, 37 (1981)

[A&W70] Agarwal, B. S., Wolf, E: Phys. Rev. **D2**, 2161, 2187, 2206 (1970)

[Ber80] Berezin, F. A.: Soviet Physics – Uspekhi, vol 23, no 11, 763 (1980)

[Bou69] Bourbaki, N.: Eléments de Mathematiques, Chapter IX, Volume VI. Hermann 1969.

[Bre68] Breiman, L.: Probability. Addison-Wesley 1968

[B&S81] Blanchard, Sirugue: J. Math. Phys. **22**, 1372 (1981)

[Cam60] Cameron, R. H.: J. Math. and Phys. **39**, 126 (1960)

[Coh66] Cohen, L: J. Math. Phys. **7**, 781 (1966)

[Coh70] Cohen, L: J. Math. Phys. **11**, 3296 (1970)

[Coh76] Cohen, L: J. Math. Phys. **17**, 597 (1976)

[C&D95] Cartier, P., DeWitt-Morette, C: J. Math. Phys. **36**, 2237 (1995)

[C&D99] Chaichian, M, Demichev, A: Path Integrals in Physics. Unpublished book 2000

[C&S89] Catrigiano, Stark: J. Math. Phys. **30**, 2785 (1989)

[DeW57] DeWitt, B.S.: Rev. Mod. Phys. **29**, 337 (1957)

[DeW72] DeWitt-Morette, C: Comm. Math. Phys. **28**, 47 (1972)

[DeW74] DeWitt-Morette, C: Comm. Math. Phys. **37**, 63 (1974)

[Dir33] Dirac, P.A.M: Physikalishce Zeitschrift der Sowjetunion **3**, 64 (1933)

[DMN79] DeWitt-Morette, C et al: Phys. Rep. **50 C**, 256 (1979)

[D&E81] DeWitt-Morette, C., Elworthy, K. D.: Physics Reports vol 77, no 3, 125 (1981)

[D&K79] Duru, I., Kleinert, H.: J. Math. Phys. (1979), Duru, I., Kleinert, H.: Phys. Letters B **84**, 30 (1979)

[Ein56] Einstein, A.: Investigations on the Theory of the Brownian motion. Dover Publishing 1956 (republished).

[Fey48] Feynman, R. P.: Rev. Mod. Phys. **20**, 484 (1948)

[F&H65] Feynman, R. P., Hibbs, A. R.: Quantum Mechanics and Path Integration, McGraw–Hill 1965

[FLM91] Fischer, W. Leschke, H: J. Phys. A, vol 92, no 5, p. 3835 (1991)

[G&J81] Glimm, J., Jaffe, A.: Quantum Physics; A Functional Integral Point of View. Springer-Verlag 1981

[G&S95] Grosche, C., Steiner, F.: J. Math. Phys. **36**, 2355 (1995)

[G&Y60] Gel'Fand, Yaglom: J. Math. Phys. **1**, 48 (1960)

[Ito66] Ito, K.: in *Proceedings of the 5th Berkeley Symposium on mathematical statistics and probability. Vol. II, part 1, p. 145* University of California Press (1966)

[K&S65] Kustaanheimo, P., Stiefel, E.: J. Reine Angew. Math. **218**, 204 (1965)

[K&S70] Kerner, F., Sutcliffe, W.: J. Math. Phys. **11**, 391 (1970)

[Kle90] Kleinert, H: Path Integrals in Quantum Mechanics, Statistics and Polymer Physics. World Scientific Publishing 1990

[Kol56] Kolmogorov, A. N.: Foundations of the Theory of Probability, New York Chelsea Publishing Company 1956

[LRT82] Langouche, F., Roekaerts, R., Tirapegui, D.: Functional Integration and Semiclassical Expansions. D. Reidel Publishing Company 1982

[Meh64] Mehta, C. L.: J. Math. Phys. **5**, 677 (1964)

[Miz75] Mizrahi, M. M.: J. Math. Phys. **16**, 2201 (1975)

[Miz76] Mizrahi, M. M.: J. Math. Phys. **17**, vol. 4 (1976)

[Miz78] Mizrahi, M. M.: J. Math. Phys. **19**, 298 (1978)

[Miz79] Mizrahi. M. M.: In *Feynman Path Integrals – Proceedings, Marseille 1978*, Springer-Verlag 1979

[Miz80] Mizrahi, M. M.: Phys. Rev. D, vol 22, no 8, p. 2096 (1980)

[Miz81] Mizrahi, M. M.: J. Math. Phys. **22** 1372 (1981)

[Mor51] Morette, C: Phys. Rev. **81**, 848 (1951)

[Nel64] Nelson, E: J. Math. Phys. **5**, 332 (1964)

[O&S73] Osterwalder, Schrader: Comm. Math. Phys. **31**, 83 (1973)

[O&S75] Osterwalder, Schrader: Comm. Math. Phys. **42**, 281 (1975)

[Pru81] Prugovecki, E: Nuovo Cimento A, vol 61A, ser 2, no 2, 85 (1981)

[Riv87] Rivers, R. J: Path Integral Methods in Quantum Field Theory, Cambridge University Press 1987

[Sch81] Schulman, L. S: Techniques and applications of path integration, Wiley 1985

[Sch62] Schweber, S: J. Math. Phys. **3**, 831 (1962)

[Str66] Stratonvich, R. L.: J. Siam Control **4**, 362 (1966)

[Tar74] Tarski, J.: in *Functional Integration and its Applications*, Oxford University Press 1975

[Tes71] Testa, F.: J. Math. Phys. **12**, 1471 (1971)

[Tru76] Truman, A.: J. Math. Phys. **17**, 1852 (1976)

[Tru77] Truman, A.: J. Math. Phys. **18**, 1499 (1977)

[Wie23] Wiener, N.: J. Math. Phys **2**, 131 (1923)

[Wie24] Wiener, N.: Proc. London. Math. Soc. **22**, 454 (1924)

[Wie30] Wiener, N.: Acta Math. **55**, 117 (1930)

[Wil91] Williams, D: Probability With Martingales, Cambridge University Press 1991

[Y&D86] Young, A., DeWitt-Morette, C: Ann. Phys. **169**, 140 (1986)

[Øks95] Øksendal, B: Stochastic Differential Equations (4th edition), Springer-Verlag 1995