Gaussian Analysis

Ami Mulligan - s1806984

Master of Mathematics with Honours
School of Mathematics
University of Edinburgh
March 2022

Abstract

Gaussian analysis is a mathematical field stemming from probability and measure theory. It extends the concept of analysis on Euclidean spaces to analysis on Gaussian spaces, through a number of analogous results.

Gaussian analysis has many applications in a number of mathematical fields. For example, it provides the basis for the field Stochastic Differential Equations, which in turn has many applications from Stochastic modelling to important financial models. In addition to this tangible application, Gaussian analysis also proves useful in fields such as Quantum Field Theory, where the idea of a Gaussian Free Field is common.

We aim to present an extensive exposition of the framework for Gaussian analysis, as well as a number of key results in the field, such as Fernique's integrability theorem, the Borel isoperimetirc inequality, and the basics of Malliavin calculus. This dissertation will serve as a collected document presenting the most important results in the study of Gaussian analysis, with discussion on the implications and further applications of these results.

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Ami Mulligan - s1806984*)

I would like to thank Ilya Chevyerev for all the help that he has provided throughout the course of this dissertation, from research to writing.

Contents

Introduction

An undergraduate mathematics student will be well versed in Euclidean spaces, and somewhat knowledgeable of analysis on Euclidean spaces. For example, calculus on Euclidean spaces is extremely useful, and plays a vital role in many different areas of mathematics. On the other hand, Gaussian spaces, and analysis thereon, is rather less familiar. This dissertation aims to introduce the framework required to do analysis on Gaussian spaces, before presenting some key results in the field of Gaussian analysis, and finally deriving a calculus for Gaussian probability spaces.

In Chapters 1 and 2 we will discuss definitions, characterisations, and examples of Gaussian measures, starting with the familiar case of a Gaussian distribution on $\mathbb{R}$, before abstracting to Gaussian measures on $\mathbb{R}^n$, and then introducing the concept of a Gaussian measure on an infinite dimensional Banach space. In addition to this we will discuss the structure of a Gaussian probability space through the Cameron-Martin theorem, and its implications.

The main focus of this dissertation is a number of key results, from Fernique's theorem on the integrability of Gaussian measures, to the Gaussian analogue of the isoperimetric inequality, and finally the development of Malliavin calculus for random variables, and the Wiener chaos decomposition of probability spaces. The proofs for some of the presented results are quite involved, but we aim to present them in a clear and coherent manner in order to make them accessible to the reader.

Chapter 1

Gaussian measures

Gaussian measures are an integral part Gaussian analysis and therefore of this dissertation. A good understanding of them is necessary for what will be presented in the later chapters. Hence, we shall spend the first chapter introducing Gaussian measures through their definitions, some basic properties and a number of examples in order to provide a strong base for the rest of our discussion.

1.1 Gaussian measures on the real line

We will begin by introducing a Gaussian measure on the real line. This is the simplest version of a Gaussian measure, and will allow us to build some intuition with an object which we are already familiar with.

Definition 1.1.1 ([1])**.** A Borel probability measure, γ on $\mathbb{R}$, is *Gaussian* if it is either the *Dirac measure* δ_a at a point, or has density

$$p(\cdot, a, \sigma^2) : x \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)$$

with respect to the *Lebesgue measure* on $\mathbb{R}$. In the latter case, γ is said to be *non-degenerate*. We call a and σ^2 the *mean* and *variance* of γ respectively.

We now introduce a proposition which allows us to completely determine a measure.

Proposition 1.1.2 ([2])**.** *If two bounded Borel measures have equal Fourier transforms, then they coincide.*

Proof. We will outline the proof of this proposition following the proof found in Proposition 3.8.6 in Bogachev's *Measure Theory* [2].

To this end, it suffices to show that if μ is a measure on $\mathbb{R}^n$ such that its Fourier transform, $\widehat{\mu}$, is identically zero, then so is μ. Suppose that μ is such that $\widehat{\mu} = 0$, then we can assume that $\|\mu\| \leq 1$ since it is bounded. Furthermore, consider a bounded continuous function f, and we may similarly assume that $|f| \leq 1$. Finally, take $\varepsilon \in (0,1)$. Now let f_0 be a continuous function with bounded support such that $|f_0| \leq 1$ and

$$\int_{\mathbb{R}^n} \left|f(x) - f_0(x)\right| \,\mathrm{d}|\mu|\,(x) \leq \varepsilon.$$

Since f_0 has bounded support, we can find a cube $K = [-\pi k, \pi k]^n$ containing the support of f_0 such that $|\mu|\big(\mathbb{R}^n \setminus K\big) < \varepsilon$. Now by the Stone-Weierstrass theorem, there exists a continuous function of the form

$$g(x) = \sum_{1 \le j \le m} c_j \exp\Big(i \langle y_j, x\rangle\Big),$$

where the y_j are vectors with coordinates of the form $\frac{l}{k}$, such that

$$\big|f_0(x) - g(x)\big| \le \varepsilon,$$

for all $x \in K$. Notice that g is periodic on $\mathbb{R}^n$ by construction, hence $\big|g(x)\big| < 1 + \varepsilon < 2$ on $\mathbb{R}^n$. Moreover, $\int_{\mathbb{R}^n} g \,\mathrm{d}\mu = 0$, since $\widehat{\mu} = 0$. Now it just remains to put all the parts together.

$$\begin{aligned}
\left|\int_{\mathbb{R}^n} f \,\mathrm{d}\mu\right| &\le \left|\int_{\mathbb{R}^n} f - f_0 \,\mathrm{d}\mu\right| + \left|\int_{\mathbb{R}^n} f_0 \,\mathrm{d}\mu\right| \\
&\le \varepsilon + \left|\int_{\mathbb{R}^n} f_0 - g + g \,\mathrm{d}\mu\right| \\
&\le 2\varepsilon + \int_{\mathbb{R}^n \setminus K} |g| \,\mathrm{d}|\mu| \\
&\le 4\varepsilon.
\end{aligned}$$

Taking $\varepsilon \to 0$ gives the result. □

With this result, it is natural to ask what the Fourier transform of an arbitrary Gaussian measure, γ on $\mathbb{R}$, looks like. To this end, we evaluate the integral

$$\widehat{\gamma}(y) := \int_{\mathbb{R}} \exp(iyx) \,\mathrm{d}\gamma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} \exp(iyx) \exp\left(-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2\right) \mathrm{d}x.$$

Using the substitution $v = \frac{x-a}{\sigma} - i\sigma y$, we obtain

$$\begin{aligned}
\widehat{\gamma}(y) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(iy(\sigma v + i\sigma^2 y + a) - \frac{1}{2}(v + i\sigma y)^2\right) \mathrm{d}v \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(iay - \frac{1}{2}\sigma^2 y^2\right) \int_{\mathbb{R}} \exp\left(-\frac{1}{2}v^2\right) \mathrm{d}v \\
&= \exp\left(iay - \frac{1}{2}\sigma^2 y^2\right).
\end{aligned}$$

We can state this formally as below.

Proposition 1.1.3. *Let γ be an arbitrary Gaussian measure on $\mathbb{R}$, then its Fourier transform is of the form*

$$\widehat{\gamma}(y) = \exp\left(iay - \frac{1}{2}\sigma^2 y^2\right).$$

Remark. Notice that this holds even when γ is the Dirac measure at a point a, since we have

$$\int_{\mathbb{R}} f(y) \,\mathrm{d}\delta_a(y) = f(a),$$

and the Dirac measure corresponds to a Gaussian measure with mean a and zero variance.

1.2 Gaussian measures on $\mathbb{R}^n$

We now extend some of the concepts from the previous section to Gaussian measures on $\mathbb{R}^n$.

Definition 1.2.1 ([1])**.** A Borel probability measure, γ on $\mathbb{R}^n$, is *Gaussian* if for any linear functional f on $\mathbb{R}^n$, the induced measure $\gamma \circ f^{-1}$ is Gaussian.

Example 1.2.2. We will now look at some examples.

(i). Consider the measure, μ on $\mathbb{R}^n$ with density with respect to the Lebesgue measure given by

$$\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}(x) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\|x\|^2}.$$

Then one can verify that for any linear functional f on $\mathbb{R}^n$, by considering f as an element of $\mathbb{R}^n$, we have

$$\widehat{\mu \circ f^{-1}}(t) = \exp\left(-\frac{1}{2}t^2\|f\|^2\right).$$

Comparing this to Proposition 1.1.3, we see that $\mu \circ f^{-1}$ is a Gaussian measure. In fact μ is the standard Gaussian measure on $\mathbb{R}^n$, denoted γ_n.

(ii). On the other hand, suppose that we have a measure ρ on $\mathbb{R}^2$ such that the density with respect to the Lebesgue measure is given by

$$\frac{\mathrm{d}\rho}{\mathrm{d}\lambda}(x) = \begin{cases} \frac{1}{\pi} e^{-\frac{1}{2}\|x\|^2} & \text{if } x \in [0,\infty)^2 \cup (-\infty, 0]^2 \\ 0 & \text{otherwise} \end{cases}.$$

Firstly notice that if we take $f = \frac{1}{\sqrt{2}}(1,1)^T$, then

$$\widehat{\rho \circ f^{-1}}(t) = \frac{1}{2} e^{-\frac{1}{2}t^2},$$

which is not of the form required to be Gaussian by Proposition 1.1.3. However, if we consider $e_1 = (1,0)^T$, then

$$\widehat{\rho \circ e_1^{-1}}(t) = e^{-\frac{1}{2}t^2},$$

which implies that $\rho \circ e_1^{-1}$ is Gaussian. Similarly, we see that for $e_2 = (0,1)^T$, we have $\rho \circ e_2^{-1}$ is Gaussian. This implies that each component has a Gaussian distribution, but their sum does not.

We now extend Proposition 1.1.3 to Gaussian measures on $\mathbb{R}^n$.

Proposition 1.2.3. *A measure γ, on $\mathbb{R}^n$ is Gaussian if and only if its Fourier transform has the form*

$$\widehat{\gamma}(y) = \exp\left(i\langle y, a\rangle - \frac{1}{2}\langle Ky, y\rangle\right), \tag{1.1}$$

where $a \in \mathbb{R}^n$ is a vector and K is a nonnegative matrix.

The measure γ has density if and only if K is nondegenerate, and in this case the density is given by

$$p(\cdot, a, K) : x \mapsto \frac{1}{\sqrt{(2\pi)^n \det K}} \exp\left(-\frac{1}{2}\left\langle K^{-1}(x-a), x-a\right\rangle\right).$$

Proof. Firstly suppose that γ is a Borel measure on $\mathbb{R}^n$ with Fourier transform given by (1.1). Then for any linear functional $f : \mathbb{R}^n \to \mathbb{R}$ we can consider the measure $\nu = \gamma \circ f^{-1}$ on a 1-dimensional subspace of $\mathbb{R}^n$. Let us consider the Fourier transform of ν,

$$\widehat{\nu}(t) = \int_{\mathbb{R}} \exp(its)\, \mathrm{d}\nu(s) = \int_{\mathbb{R}^n} \exp\big(itf(x)\big)\, \mathrm{d}\gamma(x).$$

Let us denote by f the linear functional considered as an element of $\mathbb{R}^n$. This gives

$$\begin{aligned} \widehat{\nu}(t) &= \int_{\mathbb{R}^n} \exp\big(it\langle f, x\rangle\big)\, \mathrm{d}\gamma(x) \\ &= \exp\left(it\langle f, a\rangle - \frac{1}{2}t^2\langle Kf, f\rangle\right), \end{aligned}$$

by assumption. Hence ν is a Gaussian measure by Proposition 1.1.3, and since this holds for arbitrary f, we have that γ is a Gaussian measure.

For the other direction, assume that γ is a Gaussian measure, so all measures of the form $\nu = \gamma \circ f^{-1}$ are Gaussian for a linear functional f. Let us denote their means and variances by $a(f)$ and $\sigma^2(f)$ respectively. Then we have the following

$$\begin{aligned} \widehat{\gamma}(f) &= \int_{\mathbb{R}^n} \exp\big(i\langle f, x\rangle\big)\, \mathrm{d}\gamma(x) \\ &= \int_{\mathbb{R}} \exp(it)\, \mathrm{d}\nu(t) \\ &= \exp\left(ia(f) - \frac{1}{2}\sigma^2(f)\right). \end{aligned}$$

However, we can write

$$\begin{aligned} a(f) &= \int_{\mathbb{R}} t\, \mathrm{d}\nu(t) = \int_{\mathbb{R}^n} f(x)\, \mathrm{d}\gamma(x), \\ \sigma^2(f) &= \int_{\mathbb{R}} \big(t - a(f)\big)^2\, \mathrm{d}\nu(t) = \int_{\mathbb{R}^n} \Big(f(x) - a(f)\Big)^2\, \mathrm{d}\gamma(x). \end{aligned}$$

So, we see that $f \mapsto a(f)$ is linear, and $f \mapsto \sigma^2(f)$ is a nonnegative quadratic form. In particular there exists a vector a and a nonnegative matrix K such that $a(f) = \langle a, f\rangle$ and $\sigma^2(f) = \langle Kf, f\rangle$. This yields (1.1).

The claim about densities can be reduced to a 1-dimensional case by choosing the coordinates corresponding to the eigenvectors of K, and considering the densities of a 1-dimensional Gaussian measure. □

Corollary 1.2.4. *Let γ be a Gaussian measure on $\mathbb{R}^n$ with Fourier transform given by*

(1.1). *Then*

$$a = \int_{\mathbb{R}^n} x \, \mathrm{d}\gamma(x),$$

$$\langle Ku, v \rangle = \int_{\mathbb{R}^n} \langle u, x-a \rangle \langle v, x-a \rangle \, \mathrm{d}\gamma(x) \quad \forall u, v \in \mathbb{R}^n \, .$$

The vector a is called the mean *of γ, and K is called the covariance operator.*

Remark. By comparing Proposition 1.2.3 to Example 1.2.2(i) we see that in the case of the standard Gaussian measure, γ_n, on $\mathbb{R}^n$, that $K = \mathbb{I}_n$.

1.3 Gaussian measures in infinite dimensions

In this section, we abstract once more to Gaussian measures on arbitrary Banach spaces. This will, however, require some extra definitions and notions, which we introduce below. We begin with a couple of definitions that will allow us to define an appropriate space upon which we can consider a Gaussian measure.

Definition 1.3.1 ([1])**.** Let F be a family of functions on a set X into $\mathbb{R}$ with the Borel σ-algebra. We denote by $\mathcal{E}(X, F)$ the minimal σ-field of subsets of X, with respect to which all functionals $f \in F$ are measurable.

Definition 1.3.2. A set of $\mathbb{R}$-valued functions F is said to *separate* the points of a set X, if for all $x, y \in X$, there exists $f \in F$ such that $f(x) \neq f(y)$.

With these definitions in hand, we can now define a Gaussian measure on an arbitrary Banach space.

Definition 1.3.3. (i). Let E be a linear space and F some linear space of linear functions separating E. A probability measure γ on $\mathcal{E}(E, F)$ is called Gaussian if, for any $f \in F$, the measure $\gamma \circ f^{-1}$ is Gaussian.

(ii). Let X be a Banach space with continuous dual X^*. A probability measure γ on $\mathcal{E}(X) := \mathcal{E}(X, X^*)$ is said to be Gaussian if for any $f \in X^*$, the induced measure $\gamma \circ f^{-1}$ on $\mathbb{R}$ is Gaussian. The measure is called centred if all the induced measures are centred (have mean 0).

(iii). A random vector is called Gaussian if it induces a Gaussian distribution, i.e. for $A \subset X$ we have that $\mathbb{P}(x \in A)$ defines a Gaussian measure.

We can now generalise Propositions 1.1.3 and 1.2.3 to an arbitrary Banach space.

Theorem 1.3.4 ([1])**.** *A measure γ on a Banach space is Gaussian if and only if its Fourier transform has the form*

$$\widehat{\gamma}(f) = \exp\left(iL(f) - \frac{1}{2} B(f, f) \right), \tag{1.2}$$

where $L : X^* \to \mathbb{R}$ is linear, and $B : X^* \times X^* \to \mathbb{R}$ is a symmetric bilinear nonnegative quadratic form.

The proof of this theorem is very similar to that of Proposition 1.2.3, but we shall present it nonetheless in order to see the subtleties of the argument.

Proof. Firstly suppose that γ is a measure on a Banach space X, such that the Fourier transform is given by (1.2). Then to show that γ is Gaussian, it suffices to show that for any $f \in X^*$ the induced measure $\nu = \gamma \circ f^{-1}$ is Gaussian. To do this consider the Fourier transform of ν below

$$\begin{aligned}\widehat{\nu}(t) &= \int_{\mathbb{R}} \exp(its)\,\mathrm{d}\nu(s) \\ &= \int_X \exp\big(itf(x)\big)\,\mathrm{d}\gamma(x) \\ &= \exp\left(itL(f) - \frac{1}{2}t^2 B(f,f)\right).\end{aligned}$$

Hence, by Proposition 1.1.3, ν is Gaussian, and therefore γ is Gaussian.

On the other hand, suppose that γ is a Gaussian measure, in particular for any $f \in X^*$, we have that $\nu = \gamma \circ f^{-1}$ is Gaussian on $\mathbb{R}$ with mean and variance given by $a(f)$ and $\sigma^2(f)$ respectively. Consider

$$\begin{aligned}\widehat{\gamma}(f) &= \int_X \exp\big(if(x)\big)\,\mathrm{d}\gamma(x) \\ &= \int_{\mathbb{R}} \exp(ix)\,\mathrm{d}\nu(x) \\ &= \exp\left(ia(f) - \frac{1}{2}\sigma^2(f)\right),\end{aligned}$$

where the last equality follows from Proposition 1.1.3.

Moreover, notice that

$$\begin{aligned}a(f) &= \int_{\mathbb{R}} x\,\mathrm{d}\nu(x) = \int_X f(x)\,\mathrm{d}\gamma(x), \\ \sigma^2(f) &= \int_{\mathbb{R}} \big(x - a(f)\big)^2\,\mathrm{d}\nu(x) = \int_X \big(f(x) - a(f)\big)^2\,\mathrm{d}\gamma(x).\end{aligned}$$

So $f \mapsto a(f)$ defines a linear map, and $f \mapsto \sigma^2(f)$ defines a nonnegative quadratic form, so taking $L(f) = a(f)$ and $B(f,f) = \sigma^2(f)$ gives (1.2). $\square$

In the proof of the previous theorem, it was required to show that the objects L and B existed, and in doing so, we defined $a(f)$ and $\sigma^2(f)$ to be the mean and variance of the induced measure $\nu = \gamma \circ f^{-1}$ and argued that these satisfied the requirements. In the next definition we consolidate this and formally define the mean and covariance of a Gaussian measure.

Definition 1.3.5. Let X be a Banach space and let γ be a measure on $\mathcal{E}(X)$ such that $X^* \subset L^2(\gamma)$. The element $a_\gamma \in (X^*)'$ defined by

$$a_\gamma(f) = \int_X f(x)\,\mathrm{d}\gamma(x)$$

is the *mean* of γ. If $a_\gamma(f) = 0$ for all $f \in X^*$, then we say that γ is *centred.*

The operator $R_\gamma : X^* \to (X^*)'$ defined by

$$R_\gamma(f)(g) = \int_X \Big[f(x) - a_\gamma(f)\Big]\Big[g(x) - a_\gamma(g)\Big]\,\mathrm{d}\gamma(x)$$

is the *covariance operator* of γ. The corresponding quadratic form on X^* is called the covariance of γ.

Remark. Notice in the above definition, the appearance of $(X^*)'$. Here, Y^* denotes the *continuous dual* of an arbitrary Banach space Y, while Y' denotes the *algebraic dual*, in particular an element of the algebraic dual need not be continuous.

It is worth remarking at this point that in the n-dimensional cases, we have means and covariances which are independent of $f \in (\mathbb{R}^n)^*$, whereas, in the arbitrary Banach space setting, we see a dependence on $f \in X$. This discrepancy is a result of not setting a basis in an arbitrary Banach space, whereas in $\mathbb{R}^n$ we can work with a specific basis, and so write the mean and covariance without a dependence on f.

Next we introduce a proposition that will allow us to characterise centred Gaussian measures.

Proposition 1.3.6. *Let γ be a centred Gaussian measure on the σ-algebra $\mathcal{E}(X)$ of a Banach space X. Then for every real φ, the image of $\gamma \otimes \gamma$ under the mapping*

$$\begin{array}{rcrcl} \alpha_\varphi & : & X \times X & \longrightarrow & X \\ & & (x, y) & \longmapsto & x \sin\varphi + y \cos\varphi \end{array}$$

coincides with γ.

Proof. It suffices to show that the Fourier transform of $\gamma \otimes \gamma$ under the above mapping is equal to $\widehat{\gamma}$. To this end, consider

$$\begin{aligned} \widehat{\gamma \otimes \gamma \circ \alpha_\varphi^{-1}}(f) &= \iint_{X \times X} \exp\left(if\left(x \sin\varphi + y \sin\varphi\right)\right) \, \mathrm{d}\gamma(x)\mathrm{d}\gamma(y) \\ &= \int_X \exp\left(if\left(x \sin\varphi\right)\right) \, \mathrm{d}\gamma(x) \int_X \exp\left(if\left(y \cos\varphi\right)\right) \, \mathrm{d}\gamma(y) \\ &= \int_{\mathbb{R}} \exp\left(ix \sin\varphi\right) \, \mathrm{d}\gamma \circ f^{-1}(x) \int_{\mathbb{R}} \exp\left(iy \cos\varphi\right) \, \mathrm{d}\gamma \circ f^{-1}(y) \\ &= \exp\left(i\left(\sin\varphi + \cos\varphi\right) a_\gamma(f) - \frac{1}{2}\left(\sin^2\varphi + \cos^2\varphi\right)\sigma^2(f)\right). \end{aligned}$$

Since γ is centred, $a_\gamma(f) = 0$, therefore, by Theorem 1.3.4,

$$\widehat{\gamma \otimes \gamma \circ \alpha_\varphi^{-1}}(f) = \widehat{\gamma}(f)$$

as required. □

1.3.1 Brownian motion and the Wiener measure

Here, we will introduce an instructive example of a Gaussian measure in infinite dimensions, and one that will be useful later in this dissertation. Let us begin by introducing Brownian motion.

Definition 1.3.7. A real valued stochastic process $B = \{B_t\}_{t \geq 0}$ defined on a probability space (Ω, Σ, P) is called a *Brownian motion* if it satisfies

(i). $B_0 = 0$ almost surely;

(ii). For all $0 \leq t_0 < t_1 < \cdots < t_n$, the increments $B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}}$ are mutually independent;

(iii). If $0 \leq s < t$, then $B_{t-s} \sim \mathcal{N}(0, t-s)$;

(iv). The map $t \mapsto B_t$ is continuous almost surely.

Now let us consider the space of continuous functions from the closed unit interval, $[0,1]$, to $\mathbb{R}$, $C[0,1]$, equipped with the σ-algebra $\mathscr{C}$ such that all coordinate maps $w \mapsto w(t)$, from $C[0,1]$ into $\mathbb{R}$, are measurable for all $t \in [0,1]$. Given a Brownian motion $B = \{B_t\}_{t\geq 0}$, we may consider the mapping defined by

$$\begin{aligned}(\Omega, \Sigma, P) &\to C[0,1] \\ \omega &\mapsto \big(t \mapsto B_t(\omega)\big),\end{aligned}$$

where (Ω, Σ, P) is a probability space and $B_t(\omega)$ denotes a sample path of the Brownian motion. The *Wiener measure*, P^W on $\big(C[0,1], \mathscr{C}\big)$, is then defined to be the image of P under this mapping. In particular, if we have a set $A \subset C[0,1]$, then we have

$$P^W(A) = P(B_* \in A),$$

where B_* denotes the map $\big(t \mapsto B_t(\omega)\big) \in C[0,1]$. Furthermore, it is a (non-trivial) fact that there exists a probability space (Ω, Σ, P) with a Brownian motion defined on it, and moreover, the measure P^W is independent of our choice of (Ω, Σ, P) and the Brownian motion, B (see Theorem 14.5 of Kallenberg's 'Foundations of Modern Probability' [3]).

Now suppose that $A \subset C[0,1]$ is of the form

$$A = \big\{f \in C[0,1] : f(t_0) \in A_0, \cdots, f(t_k) \in A_k\big\},$$

where $0 = t_0 < t_1 < \cdots < t_k$ and $A_0, \cdots, A_k \in \mathcal{B}(\mathbb{R})$. Then, from Definition 1.3.7, we can see that

$$\begin{aligned}P^W(A) &= P\big(B_{t_0} \in A_0, \cdots, B_{t_k} \in A_k\big) \\ &= \mathbb{1}_{A_0}(0) \int_{A_1 \times \cdots \times A_k} p_{t_1}(x_1) p_{t_2 - t_1}(x_2 - x_1) \cdots p_{t_k - t_{k-1}}(x_k - x_{k-1})\, \mathrm{d}x,\end{aligned}$$

where $p_t(x) := p(x, 0, t)$ from Definition 1.1.1 is the Gaussian density in one dimension. We then call the probability space $\big(C[0,1], \mathscr{C}, P^W\big)$ the *Wiener space.*

1.4 Notation

We end this chapter by introducing some useful notation that we will come across again later in our discussion of the Cameron-Martin space and theorem. Let γ be a Gaussian measure on a Banach space X, then we denote by X^*_γ the closure of the set

$$\big\{f - a_\gamma(f) : f \in X^*\big\}$$

embedded into $L^2(\gamma)$ with respect to the corresponding norm.

Notice that we can also define $R_\gamma : X_\gamma^* \to (X^*)'$ by having

$$R_\gamma(f)(g) = \int_X f(x) \left[g(x) - a_\gamma(g) \right] \, \mathrm{d}\gamma(x).$$

For centred measures, this is just an extension from X^* to X_γ^*. But for any $f \in X^*$, we have $R_\gamma(f)$ coincides with $R_\gamma \left(f - a_\gamma(f) \right)$.

If it is the case that $R_\gamma(f)$, with $f \in X_\gamma^*$ is generated by an element of the initial space X - for example in the n-dimensional case - then $\left\langle R_\gamma(f), l \right\rangle = \left\langle l, R_\gamma(f) \right\rangle$ for all $l \in X^*$.

Chapter 2

Cameron-Martin space

In this chapter, we would like to gain some intuition for the structure of a Banach space X with a Gaussian measure γ. An important part of this is the Cameron-Martin space, which we shall define, and then prove the Cameron-Martin theorem in addition to some related theorems at the end of the chapter.

2.1 The Cameron-Martin space

We now define the Cameron-Martin space.

Definition 2.1.1. For $h \in X$, put

$$|h|_{H(\gamma)} = \sup \left\{ l(h) : l \in X^*,\ R_\gamma(l)(l) \leq 1 \right\}.$$

We can then define the *Cameron-Martin space* to be the linear space

$$H(\gamma) = \left\{ h \in X : |h|_{H(\gamma)} < \infty \right\}.$$

This definition is rather abstract, and gives little indication as to the underlying structure of the Cameron-Martin space. We spend the rest of this chapter trying to build an understanding of this structure.

We begin by with the following lemma.

Lemma 2.1.2 ([1]). *A vector $h \in X$ belongs to the Cameron-Martin space, $H(\gamma)$ precisely when there exists $g \in X_\gamma^*$ with $h = R_\gamma(g)$. In this case we have*

$$|h|_{H(\gamma)} = \|g\|_{L^2(\gamma)}.$$

Proof. Firstly suppose that we have $|h|_{H(\gamma)} < \infty$, then by the Riesz representation theorem, there exists a $g \in X_\gamma^*$ such that for all $f \in X^*$ we have

$$f(h) = \langle f - a_\gamma(f), g \rangle = \int_X \Big(f(x) - a_\gamma(f) \Big) g(x) \, \mathrm{d}\gamma(x).$$

Therefore, we can see that $h = R_\gamma(g)$.

On the other hand, suppose that we have $h = R_\gamma(g)$ for some $g \in X^*_\gamma$. Then we have

$$\begin{aligned} |h|_{H(\gamma)} &= \sup\left\{l(h) : \ l \in X^*, \ R_\gamma(l)(l) \leq 1\right\} \\ &= \sup\left\{R_\gamma(g)(l) : \ l \in X^*, \ R_\gamma(l)(l) \leq 1\right\} \\ &= R_\gamma(g)(g) < \infty, \end{aligned}$$

hence $h \in H(\gamma)$ and $|h|_{H(\gamma)} = \|g\|_{L^2(\gamma)}$. □

Building on this, we introduce the notation $\tilde{h} := g \in X^*$ if $h = R_\gamma g$, and we say that the element $g \in X^*$ is *associated* with h. As seen in the proof of Lemma 2.1.2, we have that $\tilde{h}$ is determined by

$$f(h) = \int_X \left[f(x) - a_\gamma(f)\right] \tilde{h}(x) \, \mathrm{d}\gamma(x)$$

for all $f \in X^*$.

Let us suppose that we have two elements $\tilde{h}, \tilde{k}$ defined as above, then notice that $\tilde{h}, \tilde{k} \in X^*$, in particular, they are elements of the dual space of X. So let us consider the $L^2(\gamma)$ inner product on these elements

$$\left\langle \tilde{h}, \tilde{k} \right\rangle_{L^2(\gamma)} = \int_X \tilde{h}(x)\tilde{k}(x) \, \mathrm{d}\gamma(x).$$

Then it is easy to see that this defines an inner product on the Cameron-Martin space via

$$\langle h, k \rangle_{H(\gamma)} := \left\langle \tilde{h}, \tilde{k} \right\rangle_{L^2(\gamma)}.$$

Furthermore, since we have $R_\gamma(X^*_\gamma) \subset X$, then by Lemma 2.1.2, we have that $H(\gamma) = R_\gamma(X^*_\gamma)$. In particular, we see that the Cameron-Martin space, $H(\gamma)$, is a Hilbert space, with the norm given by

$$\|h\| = \sqrt{\left\langle \tilde{h}, \tilde{h} \right\rangle_{L^2(\gamma)}} = \sqrt{R_\gamma\left(\tilde{h}\right)\left(\tilde{h}\right)}.$$

Moreover, Lemma 2.1.2 implies that the mapping $R_\gamma : X^*_\gamma \to H(\gamma)$ is in fact an isomorphism. This allows us to think of the Cameron-Martin space in terms of centred Gaussian variables on our space.

Example 2.1.3. Before proceeding, let us look at some examples of Cameron-Martin spaces.

(i). Suppose that we have a non-degenerate Gaussian measure, γ, on $\mathbb{R}^1$, then $H(\gamma)$ is equal to $\mathbb{R}^1$, since $\alpha x < \infty$ for all $\alpha, x \in \mathbb{R}^1$.

(ii). The above example generalises to non-degenerate Gaussian measures on $\mathbb{R}^n$.

(iii). Suppose ρ is the law of (X_1, X_2) such that $X_2 \equiv 0$ and $X_1 \sim \mathcal{N}(0, 1)$. Then ρ is degenerate. Moreover we can take $h = (0, 1)^T$ and $l = (0, \alpha)^T$ with $\alpha \in \mathbb{R}$. In particular

$$R_\rho(l)(l) = \int_{\mathbb{R}^2} \alpha^2 x_2^2 \, \mathrm{d}\rho(x) = 0,$$

but $\langle l, h\rangle = \alpha$, and since α is arbitrary, we see that $|h|_{H(\rho)} = \infty$. In particular $h \notin H(\rho)$. Since $(1,0)^T$ and $(0,1)^T$ form a basis for $\mathbb{R}^2$, it follows that

$$H(\rho) = \left\{x \in \mathbb{R}^2 : x_2 = 0\right\}.$$

2.2 Cameron-Martin theorem

This section aims to provide alternate characterisations of the Cameron-Martin space in terms of equivalent measures. Let us begin by stating the theorem we wish to prove.

Theorem 2.2.1 (Cameron-Martin [4]). *Let γ be a Gaussian measure on a Banach space X.*

(i). Let $h \in X$ be a vector such that $|h|_{H(\gamma)} = \infty$, then the measures $\gamma_h := \gamma(\cdot - h)$ and γ are mutually singular.

(ii). Let $h \in X$ be a vector such that $|h|_{H(\gamma)} < \infty$, then the measures γ_h and γ are equivalent. In particular, we have the following characterisations of the Cameron-Martin space

$$H(\gamma) = \left\{h \in X : |h|_{H(\gamma)}\right\} = \{h \in X : \gamma_h \sim \gamma\} = X \cap R_\gamma\left(X_\gamma^*\right).$$

In order to prove this theorem, we will require a lemma allowing us to bound the *total variation distance*

$$\|\mu - \nu\| = \sup_{A \in \Sigma}\left|\mu(A) - \nu(A)\right|$$

between two probability measures μ, ν. Informally this can be thought of as the maximum difference that two probability measures can assign to the same element of the σ-algebra, Σ. The reason we wish to consider the total variation distance is that it gives a good way of characterising mutually singular measures, in particular, if it is the case that $\|\mu - \nu\| = 1$, then it easy to see that there exists a set, $A \in \Sigma$ upon which, without loss of generality, $\mu(A) = 1$ and $\nu(A) = 0$.

Lemma 2.2.2. *Let γ be a Gaussian measure on $\mathbb{R}^n$. Then for any vector $h \in R_\gamma(\mathbb{R}^n)$, one has the bounds*

$$1 - \exp\left(-\frac{1}{8}|h|^2_{H(\gamma)}\right) \leq \|\gamma_h - \gamma\| \leq \sqrt{1 - \exp\left(-\frac{1}{4}|h|^2_{H(\gamma)}\right)}.$$

Proof. Notice that it suffices to prove the statement for centred non-degenerate Gaussian measures on $\mathbb{R}^n$, since if γ is degenerate, then we can prove the statement on the support of γ, and for any $h \in X$ not in the support of γ, we have the equalities $|h|_{H(\gamma)} = \infty$ and $\|\gamma_h - \gamma\| = 1$. So we assume that γ is a centred non-degenerate Gaussian measure on $\mathbb{R}^n$. In this case, notice that any such measure is the image of the standard Gaussian measure, μ on $\mathbb{R}^n$, under an invertible, linear mapping T, in particular, we have $\gamma_{Th} - \gamma = (\mu_h - \mu) \circ T^{-1}$, and $|Th|_{H(\gamma)} = |h|_{H(\mu)}$. So, we may assume that γ is the standard Gaussian measure on $\mathbb{R}^n$. Moreover, we may then assume that $h = \alpha e_1$ for some $\alpha \in \mathbb{R}$. This allows us to reduce the argument to the 1-dimensional case.

Suppose that γ is the standard Gaussian measure on $\mathbb{R}^1$, then γ has density with respect to the Lebesgue measure given by

$$p = \frac{1}{2\pi} \exp\left(-\frac{1}{2}x^2\right),$$

and γ_h has density given by

$$q = \frac{1}{2\pi} \exp\left(\frac{1}{2}(x-h)^2\right).$$

Now we have that $\|\gamma_h - \gamma\| = \frac{1}{2}\|q-p\|_{L^1}$, so let us consider the inequality

$$\left(\sqrt{q}-\sqrt{p}\right)^2 \leq |q-p| = \left|\sqrt{q}-\sqrt{p}\right|\left|\sqrt{q}+\sqrt{p}\right|.$$

Taking the integral of the above inequality with respect to the Lebesgue measure gives

$$\begin{aligned}\int_{\mathbb{R}} \left(\sqrt{q}-\sqrt{p}\right)^2 \,\mathrm{d}x \leq \|q-p\|_{L^1} &= \int_{\mathbb{R}} \left|\sqrt{q}-\sqrt{p}\right|\left|\sqrt{q}+\sqrt{p}\right| \,\mathrm{d}x \\ &\leq \sqrt{\int_{\mathbb{R}} \left|\sqrt{q}-\sqrt{p}\right|^2 \,\mathrm{d}x \int_{\mathbb{R}} \left|\sqrt{q}+\sqrt{p}\right|^2 \,\mathrm{d}x},\end{aligned}$$

where the second inequality follows by Cauchy-Schwarz. Notice however, that

$$\begin{aligned}\int_{\mathbb{R}} \left(\sqrt{q}-\sqrt{p}\right)^2 \,\mathrm{d}x &= \int_{\mathbb{R}} q + p \,\mathrm{d}x - 2\int_{\mathbb{R}} \sqrt{q}\sqrt{p}\,\mathrm{d}x \\ &= 2 - 2\exp\left(-\frac{1}{8}h^2\right),\end{aligned}$$

where we note that since γ_h and γ are probability distributions we have that

$$\int_{\mathbb{R}} q \,\mathrm{d}x = \int_{\mathbb{R}} p \,\mathrm{d}x = 1,$$

and

$$\begin{aligned}\int_{\mathbb{R}} \sqrt{q}\sqrt{p}\,\mathrm{d}x &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{4}h^2\right) \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(x^2 - hx\right)\right) \,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{8}h^2\right) \int_{\mathbb{R}} \exp\left(-\frac{1}{2}u^2\right) \,\mathrm{d}u \\ &= \exp\left(-\frac{1}{8}h^2\right).\end{aligned}$$

Above the third inequality follows by the substitution $u = x - \frac{1}{2}h$.

Similarly, it can be shown that

$$\int_{\mathbb{R}} \left|\sqrt{q}+\sqrt{p}\right|^2 \,\mathrm{d}x = 2 + 2\exp\left(-\frac{1}{8}h^2\right).$$

Substituting these into the previous inequality gives

$$2 - 2\exp\left(-\frac{1}{8}h^2\right) \leq \|p - q\|_{L^1} \leq \sqrt{\left(2 - 2\exp\left(-\frac{1}{8}h^2\right)\right)\left(2 + 2\exp\left(-\frac{1}{8}h^2\right)\right)}$$
$$= 2\sqrt{1 - \exp\left(-\frac{1}{4}h^2\right)}.$$

Finally, we can rewrite this in terms of the total variation distance of the measures to give

$$1 - \exp\left(-\frac{1}{8}|h|^2_{H(\gamma)}\right) \leq \|\gamma_h - \gamma\| \leq \sqrt{1 - \exp\left(-\frac{1}{4}|h|^2_{H(\gamma)}\right)},$$

completing the proof. □

We now have all the tools to prove Theorem 2.2.1.

Proof of Theorem 2.2.1. (i). Let γ be a Gaussian measure and $h \in X$ be such that $|h|_{H(\gamma)} = \infty$. Firstly, suppose that P is a continuous, finite dimensional, linear operator. Then suppose that we have a set E in the σ-algebra corresponding to $P(X)$ such that E maximises the value $\left|(\gamma \circ P^{-1})_{Ph}(E) - \gamma \circ P^{-1}(E)\right|$, then we can consider

$$\begin{aligned}
\left|(\gamma \circ P^{-1})_{Ph}(E) - \gamma \circ P^{-1}(E)\right| &= \left|\gamma\left(P^{-1}\left(E - Ph\right)\right) - \gamma\left(P^{-1}\left(E\right)\right)\right| \\
&= \left|\gamma\left(P^{-1}(E) - h\right) - \gamma\left(P^{-1}(E)\right)\right| \\
&= \left|\gamma_h\left(P^{-1}(E)\right) - \gamma\left(P^{-1}(E)\right)\right| \\
&\leq \sup_{A \in \Sigma_X}\left|\gamma_h(A) - \gamma(A)\right|,
\end{aligned}$$

where Σ_X is the σ-algebra corresponding to X. In particular, we can see that for any continuous, finite dimensional, linear operator

$$\|\gamma_h - \gamma\| \geq \left\|\left(\gamma \circ P^{-1}\right)_{Ph} - \gamma \circ P^{-1}\right\|.$$

By assumption, we have that for all $n \in \mathbb{N}$, there exists an $f \in X^*$ such that $R_\gamma(f)(f) \leq 1$ and $f(h) > n$. Since f is a continuous finite dimensional linear operator, we have that

$$\|\gamma_h - \gamma\| \geq \left\|\left(\gamma \circ f^{-1}\right)_{f(h)} - \gamma \circ f^{-1}\right\| \geq 1 - \exp\left(-\frac{1}{8}n^2\right),$$

where the second inequality follows from Lemma 2.2.2. Now, by taking $n \to \infty$, we see that $\|\gamma_h - \gamma\| = 1$, which is equivalent to γ_h and γ being mutually singular by previous discussion.

(ii). Suppose now that $h \in X$ with $|h|_{H(\gamma)} < \infty$, then by Lemma 2.1.2 there exists $g \in X^*_\gamma$ such that $h = R_\gamma(g)$. Now consider the measure with Radon-Nikodym density given by

$$\rho_h(x) = \exp\left(g(x) - \frac{1}{2}|h|^2_{H(\gamma)}\right),$$

and denote this measure by ν. Now consider the function, for $f \in X^*$, given by

$$\phi(z) = \exp\left(ia_\gamma(f) - \frac{1}{2}|h|^2_{H(\gamma)}\right)\int_X \exp\left(i\left(f(x) - a_\gamma(f) - zg(x)\right)\right) \mathrm{d}\gamma(x),$$

in particular, notice that $\phi(i)$ coincides with the Fourier transform of ν, $\widehat{\nu}(f)$. On the other hand, since $f - a_\gamma(f) - zg \in X^*_\gamma$, in particular denoting $r = f - a_\gamma(f) - zg$, we have that $\widehat{\gamma}(r) = \exp\left(-\frac{1}{2}R_\gamma(r)(r)\right)$, which corresponds exactly to

$$\begin{aligned}\phi(z) &= \exp\left(ia_\gamma(f) - \frac{1}{2}|h|^2_{H(\gamma)}\right)\exp\left(-\frac{1}{2}\int_X \left(f - a_\gamma(f) - zg\right)^2 \mathrm{d}\gamma\right) \\ &= \exp\left(ia_\gamma(f) - \frac{1}{2}|h|^2_{H(\gamma)}\right)\exp\left(-\frac{1}{2}\sigma^2(f) + zR_\gamma(g)(f) - \frac{1}{2}\sigma^2(g)\right).\end{aligned}$$

However, by assumption, we have that $h = R_\gamma(g)$, therefore, $\sigma^2(g) = |h|^2_{H(\gamma)}$, and $R_\gamma(g)(f) = f(h)$. Therefore,

$$\widehat{\nu}(f) = \phi(i) = \exp\left(i\left(a_\gamma(f) + f(h)\right) - \frac{1}{2}\sigma^2(f)\right),$$

and noting that $a_{\gamma_h}(f) = a_\gamma(f) + f(h)$, we see that the Fourier transform of ν coincides with the Fourier transform of γ_h obtained via Theorem 1.3.4. Hence, we can conclude that γ_h is the measure with Radon-Nikodym density given by

$$\rho_h = \exp\left(g(x) - \frac{1}{2}|h|^2_{H(\gamma)}\right)$$

with respect to γ. This implies that γ, and γ_h are equivalent.

Finally, we note that

$$H(\gamma) = \left\{h \in X : |h|_{H(\gamma)} < \infty\right\} = \{h \in X : \gamma_h \sim \gamma\} = X \cap R_\gamma(X^*_\gamma),$$

where the first equality is by definition, the second is by the argument above, and the the third is by the discussion at the end of the previous section, completing the proof.

$\square$

Chapter 3

Fernique's theorem

In this section we aim to generalise the idea that Gaussian measures have *exponential tails*. This is easy to see for Gaussian measures on $\mathbb{R}^1$, for example, if we take γ_1 to be the standard Gaussian measure on $\mathbb{R}^1$, then we can consider the expectation

$$\begin{aligned}\mathbb{E}_{\gamma_1}\left[e^{\alpha x^2}\right] &= \int_{\mathbb{R}^1} e^{\alpha x^2}\,\mathrm{d}\gamma_1(x)\\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}^1} e^{\left(\alpha-\frac{1}{2}\right)x^2}\,\mathrm{d}x\\ &= \frac{1}{\sqrt{1-2\alpha}}.\end{aligned}$$

Hence, taking $\alpha \in \left(0, \frac{1}{2}\right)$ ensures that $\mathbb{E}_{\gamma_1}\left[e^{\alpha x^2}\right] < \infty$. It is relatively easy to verify that this property holds for arbitrary Gaussian measures on $\mathbb{R}^1$.

3.1 Measurable norms

In order to understand the statement of Fernique's theorem, we introduce the notion of *measurable norms* on a Banach space X. Firstly recall that a norm is defined by the following.

Definition 3.1.1. Given a vector space X, a norm is a function $q : X \to \mathbb{R}$ such that:

(i). $q(x) \geq 0$ for all $x \in X$, and $q(x) = 0$ if and only if $x = 0$;

(ii). $q(\lambda x) = |\lambda|\, q(x)$ for all $x \in X$ and $\lambda \in \mathbb{R}$;

(iii). $q(x + y) \leq q(x) + q(y)$ for all $x, y \in X$.

Remark. Recall that a Banach space is defined to be a *complete normed vector space*.

This is a familiar definition, but it is worth having on hand as we discuss the proof of Fernique's theorem. In particular, we also define a norm to be measurable in the following way.

Definition 3.1.2 ([1])**.** Let γ be a Gaussian measure on a Banach space X. A real-valued function q, measurable with respect to $\mathcal{E}(X)$, is called a *measurable norm* if there exists some $\mathcal{E}(X)$-measurable linear subspace $X_0 \subset X$ of full γ-measure such that q is a norm on X_0.

For non-degenerate, finite dimensional Gaussian measures, on $\mathbb{R}^n$ for example, this coincides with a norm on $\mathbb{R}^n$ in the regular sense. However, suppose that we have a degenerate Gaussian measure on $\mathbb{R}^2$, then its support is some linear subspace, L, isometrically isomorphic to $\mathbb{R}$. Then we need only consider a norm on L rather than on $\mathbb{R}^2$.

3.2 Fernique's theorem

To reiterate, we wish to prove that Gaussian measures on arbitrary Banach spaces have exponential tails, in particular there exists $\alpha > 0$ such that $\exp\left(\alpha q^2\right)$ is integrable with respect to the measure, where q is a norm. Formally, the statement is as follows.

Theorem 3.2.1 (Fernique's theorem [5]). *Let γ be a centred Gaussian measure on a Banach space X, and q a $\mathcal{E}(X)$-measurable norm. Then there exists $\alpha > 0$ such that*

$$\int_X \exp\left(\alpha q^2\right) \, \mathrm{d}\gamma < \infty.$$

Proof. Let $\tau, t \in \mathbb{R}_+$ be arbitrary and consider the product

$$\gamma\left(q \leq \tau\right)\gamma\left(q > t\right) = \iint_{q(x)\leq\tau;\ q(y)>t} \mathrm{d}\gamma(x)\mathrm{d}\gamma(y).$$

Write $x = \frac{1}{\sqrt{2}}\left(u - v\right)$ and $y = \frac{1}{\sqrt{2}}\left(u + v\right)$, then notice that x and y have the same form as the mapping in Proposition 1.3.6 with $\varphi = \frac{3\pi}{4}$ and $\varphi = \frac{\pi}{4}$ respectively. Furthermore, γ is centred Gaussian measure by assumption, so applying Proposition 1.3.6 gives

$$\gamma\left(q \leq \tau\right)\gamma\left(q > t\right) = \iint_{q\left(\frac{u-v}{\sqrt{2}}\right)\leq\tau;\ q\left(\frac{u+v}{\sqrt{2}}\right)>t} \mathrm{d}\gamma(u)\mathrm{d}\gamma(v).$$

Now notice that for any $u \in X$ we have

$$q(u) \geq \frac{q(u+v) - q(u-v)}{2} > \frac{t-\tau}{\sqrt{2}}$$

if $q\left(\frac{u-v}{\sqrt{2}}\right) \leq \tau$ and $q\left(\frac{u+v}{\sqrt{2}}\right) > t$. This implies that

$$\gamma(q \leq \tau)\gamma(q > t) \leq \iint_{q(u)>\frac{t-\tau}{\sqrt{2}};\ q(v)>\frac{t-\tau}{\sqrt{2}}} \mathrm{d}\gamma(u)\mathrm{d}\gamma(v) = \left[\gamma\left(q > \frac{t-\tau}{\sqrt{2}}\right)\right]^2. \tag{3.1}$$

Since q is a norm, we can choose τ such that

$$c := \gamma(q \leq \tau) > \frac{1}{2}.$$

We may assume that $c < 1$, since if $c = 1$ then we would have that

$$\int_X \exp\left(\alpha q^2\right) \, \mathrm{d}\gamma = \int_{q(x)\leq\tau} \exp\left(\alpha q^2\right) \, \mathrm{d}\gamma \leq \exp\left(\alpha\tau^2\right) < \infty.$$

Then for this τ, we can define the following

$$p_0 := \frac{\gamma(q > \tau)}{\gamma(q \leq \tau)} < 1.$$

Subsequently define the sequence $\{t_n\}_{n \in \mathbb{N}_0}$ by

$$t_n = \tau + t_{n-1}\sqrt{2},\ n \in \mathbb{N},\ t_0 = \tau.$$

Notice that we have

$$\begin{aligned} t_n &= \tau + t_{n-1}\sqrt{2} \\ &= \tau \sum_{0 \leq i \leq n} 2^{i/2} \\ &= \tau \left(\frac{2^{(n+1)/2} - 1}{\sqrt{2} - 1} \right) \\ &= \tau \left(1 + \sqrt{2}\right) \left(2^{(n+1)/2} - 1\right). \end{aligned}$$

Finally, we define the numbers p_n for $n \in \mathbb{N}$ by

$$\gamma(q > t_n) = cp_n.$$

At this point, consider

$$c^2 p_n = \gamma\left(q \leq \tau\right) \gamma\left(q > t_n\right),$$

then by applying (3.1) we get

$$\begin{aligned} c^2 p_n &\leq \left[\gamma\left(q > \frac{t_n - \tau}{\sqrt{2}} \right) \right]^2 \\ &= \left[\gamma\left(q > \frac{\tau \sum_{0 \leq i \leq n} 2^{i/2} - \tau}{\sqrt{2}} \right) \right]^2 \\ &= \left[\gamma\left(q > \tau \sum_{0 \leq i \leq n-1} 2^{i/2} \right) \right]^2 \\ &= c^2 p_{n-1}^2. \end{aligned}$$

In other words, we have that $p_n \leq p_{n-1}^2$. Applying this inequality n times gives

$$p_n \leq p_0^{2^n} = \left(\frac{1-c}{c} \right)^{2^n}.$$

In particular, we have the inequality

$$\gamma(q > t_n) \leq c \left(\frac{1-c}{c} \right)^{2^n}. \tag{3.2}$$

Now let us choose

$$\alpha = \frac{1}{24\tau^2} \log \frac{c}{1-c}.$$

Then we can begin estimating the integral

$$\begin{aligned}\int_X \exp\left(\alpha q^2\right) \, \mathrm{d}\gamma &= \int_{q\le\tau} \exp\left(\alpha q^2\right) \, \mathrm{d}\gamma + \sum_{n\ge 0} \int_{t_n<q\le t_{n+1}} \exp\left(\alpha q^2\right) \, \mathrm{d}\gamma \\ &\le c\exp\left(\alpha\tau^2\right) + \sum_{n\ge 0} \exp\left(\alpha t_{n+1}^2\right) \gamma(t_n < q \le t_{n+1}),\end{aligned}$$

where the inequality is a consequence of $\exp : \mathbb{R} \to \mathbb{R}$ being an increasing function. Furthermore, notice that $\{t_n < q \le t_{n+1}\} \subset \{t_n < q\}$, therefore, $\gamma(t_n < q \le t_{n+1}) \le \gamma(q > t_n)$. Substituting (3.2), we can bound the integral by

$$\int_X \exp\left(\alpha q^2\right) \, \mathrm{d}\gamma \le c\exp\left(\alpha\tau^2\right) + \sum_{n\ge 0} c\left(\frac{1-c}{c}\right)^{2^n} \exp\left(\alpha t_{n+1}^2\right).$$

Now since

$$t_{n+1}^2 = \tau^2\left(1+\sqrt{2}\right)^2\left(2^{(n+2)/2}-1\right)^2 \le \tau^2\left(1+\sqrt{2}\right)^2 2^{(n+2)} = 4\tau^2\left(1+\sqrt{2}\right)^2 2^n,$$

and exp is increasing, we can bound the integral by

$$\begin{aligned}\int_X \exp\left(\alpha q^2\right) \, \mathrm{d}\gamma &\le c\exp\left(\alpha\tau^2\right) + \sum_{n\ge 0} c\left(\frac{1-c}{c}\right)^{2^n} \exp\left(4\alpha\tau^2\left(1+\sqrt{2}\right)^2 2^n\right) \\ &= c\exp\left(\alpha\tau^2\right) + c\sum_{n\ge 0} \exp\left(2^n\left(\log\frac{1-c}{c} + 4\alpha\tau^2\left(1+\sqrt{2}\right)^2\right)\right) \\ &= c\exp\left(\alpha\tau^2\right) + c\sum_{n\ge 0} \exp\left(2^n \log\frac{1-c}{c}\left(1 - \frac{4\left(1+\sqrt{2}\right)^2}{24}\right)\right).\end{aligned}$$

Since $4\left(1+\sqrt{2}\right)^2 \le 24$, we can bound the integral by

$$\begin{aligned}\int_X \exp\left(\alpha q^2\right) \, \mathrm{d}\gamma &\le c\exp\left(\alpha\tau^2\right) + c\sum_{n\ge 0} \exp\left(2^n \log\frac{1-c}{c}\right) \\ &= c\exp\left(\alpha\tau^2\right) + c\sum_{n\ge 0}\left(\frac{1-c}{c}\right)^{2^n} \\ &< \infty,\end{aligned}$$

where the last inequality follows since $\frac{1-c}{c} < 1$. Hence

$$\int_X \exp\left(\alpha q^2\right) \, \mathrm{d}\gamma < \infty$$

for $\alpha = \frac{1}{24\tau^2} \log \frac{c}{1-c}$, proving the claim. $\square$

3.2.1 Application to the Wiener space

We would now like to present an application of Fernique's theorem in an infinite dimensional space. Recall the definition of a Brownian motion, $B = \{B_t\}_{t\in[0,1]}$, in Definition 1.3.7. Then we have that B is a Gaussian process on the Wiener space $\left(C[0,1], \mathscr{C}, P^W\right)$, where $\mathscr{C}$ is the σ-algebra generated by the coordinate maps $t \mapsto \omega(t)$ for $\omega \in C[0,1]$. We then have the following characterisation of B

$$\begin{array}{cccc} B_t & : & C[0,1] & \longrightarrow & \mathbb{R} \\ & & \omega & \longmapsto & \omega(t) \end{array}.$$

Now we wish to employ a theorem by Kolmogorov on the continuity of a process $X = \{X_t\}$ on $\mathbb{R}^d$ under certain conditions.

Firstly recall that for $f \in C[0,1]$ to be p-Hölder continuous means that there exists some constant $\beta \in \mathbb{R}$ such that or all $t, s \in [0,1]$, we have

$$\left|f(t) - f(s)\right| \leq \beta |t-s|^p .$$

In particular, we can define a norm $\|\cdot\|_{p\text{-Höl}}$ by

$$\|f\|_{p\text{-Höl}} := \sup_{t\neq s} \frac{\left|f(t) - f(s)\right|}{|t-s|^p}.$$

It is then easy to see that f is p-Hölder continuous if and only if $\|f\|_{p\text{-Höl}} < \infty$.

We now introduce the following theorem by Kolmogorov.

Theorem 3.2.2 (Theorem 4.23 [3])**.** *Let X be a process on $[0,1]$ with values in a complete metric space (S, ρ), such that*

$$\mathbb{E}\left[\rho\left(X_t, X_s\right)\right]^a \leq C(a,b)|t-s|^{1+b} \quad \forall t, s \in [0,1],$$

for some constants $a, b, C(a,b) > 0$. Then a version of X is Hölder continuous of order p for all $p \in \left(0, b/a\right)$.

Proof. We omit the proof of this theorem for brevity, but it can be found under Theorem 4.23 in Kallenberg's 'Foundations of Modern Probability' [3]. $\square$

Furthermore, we have that B takes values in $\left(\mathbb{R}, |\cdot - \cdot|\right)$, where $|\cdot - \cdot|$ is the absolute value metric on $\mathbb{R}$. In particular, B takes values in a complete metric space. Let $a > 0$ and let us consider the expectation for $t, s \in [0,1]$

$$\mathbb{E}|B_t - B_s|^a = C(a)|t-s|^{a/2},$$

where $C(a)$ is a constant dependant only on a. So by choosing $a > 2$ and $b = \frac{a-2}{2}$, we have that

$$\mathbb{E}|B_t - B_s|^a = C(a)|t-s|^{1+b} \quad \forall t, s \in [0,1].$$

In other words, the Brownian motion B satisfies the conditions of Theorem 3.2.2. Moreover, by the definition of a Brownian motion, the paths are almost surely continuous, so we may drop the 'a version of' condition in the result of Theorem 3.2.2. In particular, we conclude that B is p-Hölder continuous on $[0,1]$ for $p \in \left(0, b/a\right) = \left(0, \frac{1}{2} - \frac{1}{a}\right)$. Since a is arbitrary, we get that B is p-Hölder continuous for $p \in \left(0, \frac{1}{2}\right)$.

Returning to our previous discussion, we have that B is p-Hölder continuous or all $p \in \left(0, \frac{1}{2}\right)$, in other words $\|B\|_{p\text{-Höl}} < \infty$ almost surely for $p \in \left(0, \frac{1}{2}\right)$. This is equivalent to saying that

$$P^W \left(\|B\|_{p\text{-Höl}} < \infty \right) = 1,$$

in other words P^W gives full measure to the subspace of $C\,[0,1]$ composed of p-Hölder continuous functions, denoted $C^{p\text{-Höl}} \subset C\,[0,1]$. Furthermore, $\|\cdot\|_{p\text{-Höl}}$ is a norm on $C^{p\text{-Höl}}$. In the language of Fernique's theorem, $\|\cdot\|_{p\text{-Höl}}$ is a measurable norm on the Wiener space. Therefore, by Theorem 3.2.1, there exists $\alpha > 0$ such that

$$\mathbb{E}\left(e^{\alpha \|B\|^2_{p\text{-Höl}}} \right) < \infty.$$

This is a very strong statement about the integrability of functions against the Wiener measure.

Chapter 4

Gaussian isoperimetric inequality

In this chapter we shall find an inequality for Gaussian spaces analogous to the isoperimetric inequality on Euclidean spaces. It turns out that the subspaces that maximise area for a fixed surface measure in a Gaussian space is a half space. Our proof will employ a geometric method, using Gaussian k-symmetrizations, which we shall introduce next.

4.1 Gaussian symmetrization

Throughout this section, we consider the probability space $\big(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \gamma_n\big)$, where γ_n is the standard Gaussian distribution on $\mathbb{R}^n$.

Definition 4.1.1. Let $1 \leq k \leq n$, L be a linear subspace of $\mathbb{R}^n$ of dimension $n-k$, and $e \in L^\perp$ be a unit vector. We define the Gaussian k-symmetrization to be the mapping $S(L,e)$ on closed subsets $A \subset \mathbb{R}^n$ as follows

$$S(L,e)(A) := \bigcup_{x \in L} \left(\big\{ y : \langle y, e \rangle \geq r \big\} \cap \Big(x + L^\perp \Big) \right),$$

where r is such that

$$\gamma_k \left(\big\{ y : \langle y, e \rangle \geq r \big\} \cap \Big(x + L^\perp \Big) \right) = \gamma_k \left(A \cap \Big(x + L^\perp \Big) \right),$$

and γ_k is the standard k-dimensional Gaussian measure the k-dimensional affine subspace $x + L^\perp$.

This is quite an abstract definition, but it is quite easy to picture in the case that $n = 1$. For example, take $n = 1$, $L = \{0\}$, and A a closed interval in $\mathbb{R}$. Then setting $\alpha = \gamma_1(A)$, we have $S(L,e)(A) = \big(\Phi^{-1}(1-\alpha), \infty\big)$, where $\Phi(x) = \gamma_1\left(-\infty, x\right)$, see Figure 4.1.

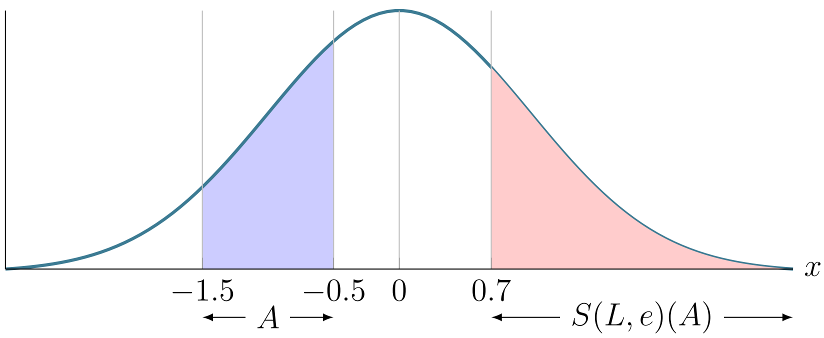

Figure 4.1: Gaussian symmetrization in the case that $n=1$, $L=\{0\}$, $A=[-1.5,-0.5]$. $\gamma_1(A)$ is given by the blue area under the density curve of γ_1. Then $S(L,e)(A)=(r,\infty)$, where $r=\Phi^{-1}\left(1-\gamma_1(A)\right)=0.7007$. Importantly the red area is equal to the blue area.

This idea can be extended to $\mathbb{R}^n$ in the sense that a Gaussian n-symmetrization, $S(L,e)(A)$, will give a halfspace with γ_n measure equal to $\gamma_n(A)$.

We then have the following properties of Gaussian symmetrizations.

Proposition 4.1.2. *For arbitrary closed sets $A, B \in \mathbb{R}^n$, one has:*

(i). If $A \subset B$, then $S(L,e)(A) \subset S(L,e)(B)$;

(ii). $S(L,e)(A^C) = S(L,-e)(A)^C$;

(iii). For all $v \in L$ we have $S(L,e)(A+v) = S(L,e)(A)+v$;

(iv). If $\{A_i\}$ is an increasing sequence of open sets such that $A=\bigcup_i A_i$, then

$$S(L,e)(A)=\bigcup_i S(L,e)(A_i).$$

Proof. (i). Suppose that $A \subset B$, and fix $x \in L$, then notice that $A \cap \left(x+L^\perp\right) \subset B \cap \left(x+L^\perp\right)$. In particular, the r_B appearing in the definition $S(L,e)(B)$ is larger than or equal to the corresponding r_A appearing in the definition of $S(L,e)(A)$. This implies that $S(L,e)(A) \cap \left(x+L^\perp\right) \subset S(L,e)(B) \cap \left(x+L^\perp\right)$, and then taking the union over all $x \in L$ gives the result.

(ii). Notice for a fixed $x \in L$ that since $\gamma_k\left(A \cap \left(x+L^\perp\right)\right)+\gamma_k\left(A^C \cap \left(x+L^\perp\right)\right)=1$, therefore, we have that $S(L,-e)(A) \cap \left(x+L^\perp\right)$ and $S(L,e)(A^C) \cap \left(x+L^\perp\right)$ are halfspaces of $x+L^\perp$ in opposite directions, such that the sum of their γ_k measures is 1. This implies that $\left(S(L,-e)(A) \cap \left(x+L^\perp\right)\right)^C = S(L,e)(A^C) \cap \left(x+L^\perp\right)$. Again, taking the union over all $x \in L$ gives the result.

(iii). Since for fixed $v \in L$ and fixed $x \in L$, we have that

$$\gamma_k\left(A \cap \left(x+L^\perp\right)\right)=\gamma_k\left((A+v) \cap \left(x+v+L^\perp\right)\right),$$

therefore, we have

$$S(L,e)(A) \cap \left(x+L^\perp\right)+v=S(L,e)(A+v) \cap \left(x+v+L^\perp\right),$$

and taking the union over all $x \in L$ gives the result.

(iv). For fixed $x \in L$, since we have that $\{A_i\}$ is an increasing sequence we know that

$$\gamma_k\left(\bigcup_i A_i \cap \left(x + L^\perp\right)\right) = \lim_{i\to\infty} \gamma_k\left(A_i \cap \left(x + L^\perp\right)\right).$$

This implies that

$$S(L,e)\left(\bigcup_i A_i\right) \cap \left(x + L^\perp\right) = \bigcup_i S(L,e)(A_i) \cap \left(x + L^\perp\right),$$

and taking the union over all $x \in L$ gives the result.

□

4.2 The Gaussian isoperimetric inequality

There are a number of proofs of the Gaussian isoperimetric inequality, in particular, in the 1970's both Borell [6] and Sudakov [7] were able to prove the isoperimetric inequality. In the 1980's Ehrhard was able to prove the inequality using Gaussian symmetrizations [8]. In the same paper, Ehrhard derived an inequality similar to the Brunn-Minkowski inequality in Euclidean space, which offers another proof of the Gaussian isoperimetric inequality [9]. It is the latter argument that we shall discuss in this dissertation.

We first introduce Ehrhard's inequality.

Theorem 4.2.1 (Ehrhard's inequality)**.** *Let A and B be two Borel sets in $\mathbb{R}^n$. Then for all $\lambda \in [0,1]$*

$$\Phi^{-1}\left\{\gamma_n\left(\lambda A + (1-\lambda)B\right)\right\} \geq \lambda\Phi^{-1}\left\{\gamma_n(A)\right\} + (1-\lambda)\Phi^{-1}\left\{\gamma_n(B)\right\}. \tag{4.1}$$

In the statement of the original theorem by Antoine Ehrhard [8], we require that $A, B \subset \mathbb{R}^n$ are convex sets, however, above we state the inequality for any measurable sets A, B in $\mathbb{R}^n$. The proof of this, however, is outside of the scope of this dissertation, but can be found in Borell's 2003 paper 'The Ehrhard inequality' [10]. Below, we present the proof under the assumption that A and B are convex sets.

Proof. First, let us consider $A + e_{n+1}$ and B to be sets in $\mathbb{R}^{n+1}$, where $\{e_i\}_{1\leq i\leq n+1}$ is the standard basis for $\mathbb{R}^{n+1}$. Then take C to be the convex hull of $A + e_{n+1}$ and B, then we may define

$$C_\lambda := \mathbb{R}^n \cap (C - \lambda e_{n+1}) = \lambda A + (1-\lambda)B.$$

We now define the function

$$f(\lambda) = \Phi^{-1}\left\{\gamma_n(C_\lambda)\right\}.$$

In particular notice that (4.1) is equivalent to

$$f(\lambda) \geq \lambda f(1) + (1-\lambda)f(0). \tag{4.2}$$

To show that (4.2) holds, let us consider the Gaussian n-symmetrization $S(L,e)$ with $L = \mathbb{R}^1 e_{n+1}$ and $e \in \{e_i\}_{1\le i\le n}$. Then we have that $r_\lambda = -f(\lambda)$ to be the r appearing in the definition of Gaussian symmetrizations. Notice that $S(L,e)(C)$ is convex, therefore

$$\lambda\left((e_{n+1}+\mathbb{R}^n)\cap S(L,e)(C)\right) + (1-\lambda)\left(\mathbb{R}^n \cap S(L,e)(C)\right) \subset (\lambda e_{n+1}+\mathbb{R}^n)\cap S(L,e)(C).$$

Specifically, this implies that $\lambda\,(r_1,\infty) + (1-\lambda)\,(r_0,\infty) \subset (r_\lambda,\infty)$, or equivalently $r_\lambda \le \lambda r_1 + (1-\lambda)r_0$. This in turn implies that (4.2) holds, thus completing the proof. □

Finally, we arrive at the Gaussian isoperimetric inequality.

Theorem 4.2.2. *Let γ_n be the standard Gaussian measure on $\mathbb{R}^n$, and let U be the closed unit ball in $\mathbb{R}^n$. Then for any measurable set $A \in \mathbb{R}^n$, and any $\varepsilon \ge 0$, the following holds*

$$\Phi^{-1}\left\{\gamma_n\left(A+\varepsilon U\right)\right\} \ge \Phi^{-1}\left\{\gamma_n\left(A\right)\right\} + \varepsilon. \tag{4.3}$$

Proof. This statement follows almost immediately from Theorem 4.2.1. Suppose that A is a measurable set in $\mathbb{R}^n$, then we consider the following

$$\begin{aligned}\Phi^{-1}\left\{\gamma_n\left(A+\varepsilon U\right)\right\} &= \Phi^{-1}\left\{\gamma_n\left(\lambda(\lambda^{-1}A) + (1-\lambda)(1-\lambda)^{-1}\varepsilon U\right)\right\}\\ &\ge \lambda\Phi^{-1}\left\{\gamma_n\left(\lambda^{-1}A\right)\right\} + (1-\lambda)\Phi^{-1}\left\{\gamma_n\left((1-\lambda)^{-1}\varepsilon U\right)\right\}.\end{aligned}$$

Taking $\lambda \to 1$, the right hand side goes to $\Phi^{-1}\left\{\gamma_n(A)\right\} + \varepsilon$, giving (4.3). □

It is not exactly obvious how Theorem 4.2.2 is related to an isoperimetric inequality, or indeed, what set maximises the volume to surface area ratio under the standard Gaussian measure. In order to understand this inequality, let us rewrite it as

$$\gamma_n\left(A+\varepsilon U\right) \ge \Phi\left(a+\varepsilon\right), \tag{4.4}$$

where $a = \Phi^{-1}\left\{\gamma_n(A)\right\}$. However, notice that $\Phi\left(a+\varepsilon\right)$ is then equal to $\gamma_n\left(H+\varepsilon U\right)$, where H is a halfspace such that $\gamma_n(H) = \gamma_n(A)$. Such an H exists since γ_n is a bounded measure for any n. Therefore, we can rewrite (4.4) as

$$\gamma_n\left(A+\varepsilon U\right) \ge \gamma_n\left(H+\varepsilon U\right). \tag{4.5}$$

We may then define the *surface measure*, γ_n^{surf}, of a set $A \in \mathbb{R}^n$ as the following limit

$$\gamma_n^{\text{surf}}(A) = \lim_{\varepsilon\to 0}\frac{\gamma_n(A+\varepsilon U) - \gamma_n(A)}{\varepsilon}.$$

Intuitively, this corresponds to how much the measure of the set A changes by adding εU, which corresponds to the measure of the surface of the set. It is then evident from (4.5) that for arbitrary $A \in \mathbb{R}^n$ and a halfspace $H \in \mathbb{R}^n$ such that $\gamma_n(A) = \gamma_n(H)$, we have

$$\gamma_n^{\text{surf}}(A) = \lim_{\varepsilon\to 0}\frac{\gamma_n\left(A+\varepsilon U\right) - \gamma_n(A)}{\varepsilon} \ge \lim_{\varepsilon\to 0}\frac{\gamma_n\left(H+\varepsilon U\right) - \gamma_n(H)}{\varepsilon} = \gamma_n^{\text{surf}}(H).$$

In other words, halfspaces minimise surface measure. One can compare this to the isoperimetric inequality in Euclidean space which states that spheres minimise surface measure.

Chapter 5

Malliavin calculus

In this chapter, we will derive a 'calculus' for random variables defined on a Gaussian probability space, so called *Malliavin calculus*. This is a tool particularly useful in the analysis of stochastic differential equations, or Quantum Field Theory. We will introduce Malliavin calculus, and explore a number of theorems and applications.

5.1 Gaussian analysis on $\mathbb{R}$

As was the case when first introducing Gaussian measures, Malliavin calculus will be discussed first in the 1-dimensional setting, so that we can build an intuition before introducing it for arbitrary Gaussian measures. In this spirit, let us consider the probability space $\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \gamma_1\big)$, where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra on $\mathbb{R}$, and γ_1 is the standard Gaussian measure on $\mathbb{R}$. Now let us introduce two differential operators on $C^1(\mathbb{R})$ as follows

- *Derivative operator*: for $f \in C^1(\mathbb{R})$ define

$$\mathscr{D} f(x) := f'(x).$$

- *Divergence operator*: for $f \in C^1(\mathbb{R})$ define

$$\delta f(x) := x f(x) - f'(x).$$

Firstly notice that if we have $f, g \in C^1_p(\mathbb{R})$, and we consider the $L^2(\gamma_1)$ inner product, then we have the following

$$\begin{aligned}
\langle \mathscr{D} f, g \rangle_{L^2(\gamma_1)} &= \int_{\mathbb{R}} f'(x) g(x) \,\mathrm{d}\gamma_1(x) \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f'(x) g(x) e^{-\frac{1}{2}x^2} \,\mathrm{d}x \\
&= \int_{\mathbb{R}} f(x) \big[g(x) - x g'(x) \big] \,\mathrm{d}\gamma_1(x) \\
&= \langle f, \delta g \rangle_{L^2(\gamma_1)},
\end{aligned}$$

where the third equality is a simple application of integration by parts and noticing that both f and g are bounded by assumption. In particular we have the following lemma.

Lemma 5.1.1. *The operators $\mathscr{D}$ and δ are adjoint on $C_p^1(\mathbb{R})$ with respect to the measure γ_1.*

We also have the following lemma regarding the commutativity of the operators.

Lemma 5.1.2. *Let $f \in C^n(\mathbb{R})$, then*

$$(\mathscr{D}\delta^n - \delta^n \mathscr{D}) f = n\delta^{n-1} f.$$

Proof. The proof follows by induction. First notice that for $k = 1$, we have

$$\begin{aligned}
\mathscr{D}\delta f &= \mathscr{D}\left[xf(x) - f'(x)\right] = f(x) + xf(x) - f''(x);\\
\delta\mathscr{D} f &= \delta f'(x) = xf(x) - f''(x).
\end{aligned}$$

It follows that $(\mathscr{D}\delta - \delta\mathscr{D}) f = f$. For the inductive step, suppose that this relation holds for $1 \leq k \leq n$, then we may consider

$$\begin{aligned}
\mathscr{D}\delta^n f &= \mathscr{D}\delta\left(\delta^{n-1} f\right)\\
&= \delta\mathscr{D}\left(\delta^{n-1} f\right) + \delta^{n-1} f\\
&= \delta\left(\delta^{n-1}\mathscr{D} f + (n-1)\delta^{n-2} f\right) + \delta^{n-1} f\\
&= \delta^n \mathscr{D} f + n\delta^{n-1} f,
\end{aligned}$$

giving the desired result. □

We now introduce a system of polynomials with some very helpful properties for what we would like to achieve in this chapter. These are the Hermite polynomials defined formally below.

Definition 5.1.3. Let $H_0 = 1$, then we define the n-th *Hermite polynomial* by

$$H_n := \delta^n 1.$$

Firstly, it is worth verifying that these are indeed polynomials. For example, let us consider some small cases of n, then we have

$$\begin{aligned}
H_1(x) &= \delta 1(x) = x,\\
H_2(x) &= \delta^2 1(x) = \delta\,(x) = x^2 - 1,\\
H_3(x) &= \delta^3 1(x) = \delta\left(x^2 - 1\right) = x^3 - 3x.
\end{aligned}$$

It is therefore easy to see that Hermite polynomials are aptly named, moreover, we notice that the leading coefficient of each polynomial is 1. Secondly, notice that we have the following equality

$$H_n' = \mathscr{D}\delta^n 1 = \delta^n \mathscr{D} 1 + n\delta^{n-1} 1 = nH_{n-1}.$$

In particular this gives the following recursion relation for Hermite polynomials

$$H_{n+1}(x) = \delta H_n(x) = xH_n(x) - H_n'(x) = xH_n(x) - nH_{n-1}(x).$$

The next proposition motivates our interest in Hermite polynomials.

Proposition 5.1.4. *The sequence of normalised Hermite polynomials $\left\{\frac{1}{\sqrt{n!}} H_n\right\}_{n\geq 0}$ defines a complete orthonormal basis of $L^2(\gamma_1)$.*

Proof. For orthonormality, consider the inner product

$$\langle H_n, H_m \rangle_{L^2(\gamma_1)} = \langle H_n, \delta^m 1 \rangle_{L^2(\gamma_1)} = \left\langle \mathscr{D} H_n, \delta^{m-1} 1 \right\rangle_{L^2(\gamma_1)} = n \langle H_{n-1}, H_{m-1} \rangle_{L^2(\gamma_1)}.$$

In particular, by repeating the above, we have that if $n = m$, then

$$\langle H_n, H_n \rangle_{L^2(\gamma_1)} = n! \langle 1, 1 \rangle_{L^2(\gamma_1)} = n!.$$

On the other hand, without loss of generality, suppose that $n > m$, then

$$\langle H_n, H_m \rangle_{L^2(\gamma_1)} = \frac{n!}{m!} \langle H_{n-m}, 1 \rangle_{L^2(\gamma_1)} = 0.$$

For completeness, suppose that $f \in L^2(\gamma_1)$ is such that $\langle f, H_n \rangle_{L^2(\gamma_1)} = 0$ for all $n \in \mathbb{N}_0$. Then since the leading coefficient of each H_n is 1, we have that $\langle f, x^n \rangle_{L^2(\gamma_1)} = 0$ for all $n \in \mathbb{N}_0$. Therefore, considering the fourier transform of f, we have

$$\widehat{f}(t) = \int_{\mathbb{R}} f(x) e^{itx} \, \mathrm{d}\gamma_1(x) = \sum_{n \geq 0} \frac{(it)^n}{n!} \int_{\mathbb{R}} f(x) x^n \, \mathrm{d}\gamma_1(x) = 0.$$

By Lebesgue's dominated convergence, the swapping of the sum and the integral is justified. Therefore, we have that $\widehat{f}(t) = 0$, implying that $f = 0$. □

The immediate implication of Proposition 5.1.4 is that given an $f \in L^2(\gamma_1)$, we can decompose it as below

$$f = \sum_{n \geq 0} \frac{1}{n!} \langle f, H_n \rangle_{L^2(\gamma_1)} H_n.$$

Specifically, we can write the function e^{ax} as follows

$$\begin{aligned} e^{ax} &= \sum_{n \geq 0} \frac{1}{n!} \langle e^{a\cdot}, H_n \rangle_{L^2(\gamma_1)} H_n(x) \\ &= \sum_{n \geq 0} \frac{a^n}{n!} \langle e^{a\cdot}, 1 \rangle_{L^2(\gamma_1)} H_n(x) \end{aligned}$$

since δ and $\mathscr{D}$ are adjoint. Finally, noting that $\langle e^{a\cdot}, 1 \rangle_{L^2(\gamma_1)} = e^{\frac{1}{2}a^2}$, we have that

$$e^{ax - \frac{1}{2}a^2} = \sum_{n \geq 0} \frac{a^n}{n!} H_n(x). \tag{5.1}$$

5.1.1 Ornstein-Uhlenbeck operator

We now introduce the *Ornstein-Uhlenbeck operator*, L, a second order differential operator defined via

$$Lf := -\delta \mathscr{D} f = -x f'(x) + f''(x). \tag{5.2}$$

At this point, we really begin to see the value of Hermite polynomials. In particular, let us consider applying the Ornstein-Uhlenbeck operator to H_n as follows

$$LH_n = -\delta \mathscr{D} H_n = -n \delta H_{n-1} = -n H_n.$$

In other words, H_n is an eigenvector of L with eigenvalue $-n$.

Now let us introduce the the operators $\{P_t\}_{t\geq 0}$ defined by $P_t H_n = e^{-nt} H_n$. Then it is easy to see that this defines a semigroup with the operation $P_t P_s = P_{t+s}$, in fact, this is called the *Ornstein-Uhlenbeck semigroup*. Suppose we have a function $f \in L^2(\gamma_1)$, then the Ornstein-Uhlenbeck semigroup acts on f by

$$P_t f = \sum_{n\geq 0} \frac{1}{n!} \langle f, H_n \rangle_{L^2(\gamma_1)} e^{-nt} H_n.$$

Moreover, we have that

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t f = \sum_{n\geq 0} \frac{1}{n!} \langle f, H_n \rangle_{L^2(\gamma_1)} \left[-ne^{-nt} H_n\right] = \sum_{n\geq 0} \frac{1}{n!} \langle f, H_n \rangle_{L^2(\gamma_1)} L\left(e^{-nt} H_n\right) = LP_t f,$$

so we see that L is the generator of the Ornstein-Uhlenbeck semigroup, justifying the name.

The following gives us an equivalent characterisation of the Ornstein-Uhlenbeck semigroup, and arguably a more familiar one.

Proposition 5.1.5. *For any $f \in L^2(\gamma_1)$, the following holds*

$$P_t f(x) = \int_{\mathbb{R}} f\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right) \mathrm{d}\gamma_1(y) = \mathbb{E}_{\gamma_1}\left[e^{-t}x + \sqrt{1-e^{-2t}}Y\right],$$

where $Y \sim \mathcal{N}(0,1)$.

Proof. Since $\{e^{ax}\}_{a\in\mathbb{R}}$ is a complete system in $L^2(\gamma_1)$, it suffices to prove the claim on these functions. To this end, consider

$$\begin{aligned}\int_{\mathbb{R}} e^{axe^{-t}+a\sqrt{1-e^{2t}}y} \mathrm{d}\gamma_1(y) &= e^{axe^{-t}+\frac{1}{2}a^2\left(1-e^{-2t}\right)} \\ &= e^{\frac{1}{2}a^2} e^{ae^{-t}x-\frac{1}{2}\left(ae^{-t}\right)^2}.\end{aligned}$$

Now by (5.1), we see that

$$\begin{aligned}\int_{\mathbb{R}} e^{axe^{-t}+a\sqrt{1-e^{2t}}y} \mathrm{d}\gamma_1(y) &= e^{\frac{1}{2}a^2} \sum_{n\geq 0} \frac{a^n e^{-nt}}{n!} H_n(x) \\ &= e^{\frac{1}{2}a^2} P_t\left(e^{ax-\frac{1}{2}a^2}\right) \\ &= P_t e^{ax}.\end{aligned}$$

So indeed the two operators coincide on $\{e^{ax}\}_{a\in\mathbb{R}}$, hence they coincide on $L^2(\gamma_1)$. □

5.2 Gaussian analysis on $\mathbb{R}^n$

Similarly to the 1-dimensional case, let us consider the probability space $\big(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \gamma_n\big)$, where $\mathcal{B}(\mathbb{R}^n)$ is the Borel σ-algebra on $\mathbb{R}^n$, and γ_n is the standard Gaussian measure on $\mathbb{R}^n$. We then generalise the derivative and divergence operators from the previous section as follows:

- For $f : \mathbb{R}^n \to \mathbb{R}$, we define the derivative operator, $\mathscr{D}$, by

$$\mathscr{D} f := \nabla f = \left(\tfrac{\partial f}{\partial x_1}, \cdots, \tfrac{\partial f}{\partial x_n} \right).$$

- For $u : \mathbb{R}^n \to \mathbb{R}^n$, we define the divergence operator, δ, by

$$\delta u := \sum_{1 \leq i \leq n} u_i x_i - \tfrac{\partial u_i}{\partial x_i} = \langle u, x \rangle - \nabla \cdot u,$$

  where the inner product is the standard inner product on $\mathbb{R}^n$.

It is easy to see that the previous definitions are just special cases of the new definitions. Indeed, the similarities go further in the following proposition.

Proposition 5.2.1. *The operators $\mathscr{D}$ and δ are adjoint. That is to say that for $f : \mathbb{R}^n \to \mathbb{R}$ and $u : \mathbb{R}^n \to \mathbb{R}^n$, we have*

$$\mathbb{E}\left(\langle \mathscr{D} f, u \rangle \right) = \mathbb{E}\left(f \delta u \right).$$

Proof. In this proof, first notice that the density of γ_n with respect to the Lebesgue measure on $\mathbb{R}^n$ is given by

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \|x\|^2 \right).$$

It is then easy to check that

$$\frac{\partial p}{\partial x_i}(x) = -x_i p(x).$$

Let us then consider the expectation

$$\begin{aligned}
\mathbb{E}\left(\langle \mathscr{D} f, u \rangle \right) &= \int_{\mathbb{R}^n} \langle \mathscr{D} f, u \rangle \, \mathrm{d}\gamma_n \\
&= \sum_{1 \leq i \leq n} \int_{\mathbb{R}^n} \tfrac{\partial f}{\partial x_i}(x) u_i(x) p(x) \, \mathrm{d}x \\
&= \sum_{1 \leq i \leq n} \int_{\mathbb{R}^n} f(x) \left[u_i(x) x_i - \tfrac{\partial u_i}{\partial x_i}(x) \right] p(x) \, \mathrm{d}x \\
&= \int_{\mathbb{R}^n} f \delta u \, \mathrm{d}\gamma_n \\
&= \mathbb{E}\left(f \delta u \right).
\end{aligned}$$

□

5.3 Gaussian analysis on the Wiener space

Until this point, we have only considered real-valued functions defined on $\mathbb{R}^n$ with the standard Gaussian measure γ_n, and as such, much of what has been done has been introducing new concepts in a familiar setting. Our goal is to develop a calculus for random variables, and this will require a more abstract setting, which shall be introduced below.

5.3.1 Brownian motion and the Wiener integral

Recall the definition of Brownian motion.

Definition 5.3.1. A real valued stochastic process $B = \{B_t\}_{t \geq 0}$ defined on defined on a probability space (Ω, Σ, P) is called a *Brownian motion* if it satisfies

(i). $B_0 = 0$ almost surely;

(ii). For all $0 \leq t_0 < t_1 < \cdots < t_n$, the increments $B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}}$ are mutually independent;

(iii). If $0 \leq s < t$, then $B_{t-s} \sim \mathcal{N}(0, t-s)$;

(iv). The map $t \mapsto B_t$ is continuous almost surely.

Now let us consider the space $H = L^2\left([0,1], \mathbb{R}\right)$, the continuous, real-valued, square integrable functions from $[0,1]$ into $\mathbb{R}$. For this discussion, any separable Hilbert space will suffice, however, the H we have defined will be useful later, and it is more constructive to work with a concrete example. Furthermore, suppose that $\{e_n\}_{n\in\mathbb{N}}$ is a orthonormal basis for H, so that for $h \in H$, we may write

$$h = \sum_{n\in\mathbb{N}} h_n e_n,$$

where $h_n = \langle h, e_n \rangle_H$. Now let us define a linear map $W : H \to L^2(\Omega, P)$, where (Ω, P) is a probability space such that $W(e_n) = \xi_n$, where ξ_n is a standard Gaussian random variable, in other words ξ_n can be thought of as being distributed as $\xi_n \sim \mathcal{N}(0,1)$. This map is called *Gaussian white noise*. Then for arbitrary $h \in H$, we have

$$W(h) = \sum_{n\in\mathbb{N}} h_n \xi_n.$$

It is then clear that

$$\mathbb{E}W(h) = 0, \quad \text{and} \quad \mathbb{E}W(h)W(g) = \langle h, g \rangle_H.$$

Let us now consider the function $\mathbb{1}_{[0,t)} \in H$ (this is where our specific choice of H is important), and define $B_t := W\left(\mathbb{1}_{[0,t)}\right)$. Now notice that $\mathbb{E}B_t = 0$, and moreover,

$$\begin{aligned}
\mathbb{E}\left[B_t B_s\right] &= \mathbb{E}\left[W\left(\mathbb{1}_{[0,t)}\right) W\left(\mathbb{1}_{[0,s)}\right)\right] \\
&= \left\langle \mathbb{1}_{[0,t)}, \mathbb{1}_{[0,s)} \right\rangle_H \\
&= t \wedge s.
\end{aligned}$$

It follows that $\{B_t\}_{t\geq 0}$ is a Brownian motion.

Now let us consider the span $\left\{\mathbb{1}_{[0,t)}\right\}_{t\geq 0}$. Then it is easy to see that this is the set of step functions from $[0,1]$ into $\mathbb{R}$, in other words, it is the set of functions of the form

$$\varphi_t = \sum_{0\leq k\leq n-1} \alpha_k \, \mathbb{1}_{\left[t_k, t_{k+1}\right)}(t),$$

where $0 \leq t_0 < t_1 < \cdots < t_n$ is a partition of some interval, and $\{\alpha_k\}_{0\leq k\leq n-1} \subset \mathbb{R}$. We can then define the *Wiener integral* to be

$$\int_{[0,1]} \varphi_t \,\mathrm{d}B_t := \sum_{0\leq k\leq n-1} \alpha_k \left(B_{t_{k+1}} - B_{t_k}\right).$$

By linearity, it is easy to see that

$$\int_{[0,1]} \varphi_t \,\mathrm{d}B_t = W(\varphi).$$

At this point, let us remark on some properties of the Wiener integral. Firstly, it has an expectation of 0, clearly

$$\mathbb{E}\left[\int_{[0,1]} \varphi_t \,\mathrm{d}B_t\right] = \mathbb{E}\left[W(\varphi)\right] = 0.$$

Furthermore, we can consider the second moment

$$\mathbb{E}\left[\int_{[0,1]} \varphi_t \,\mathrm{d}B_t\right]^2 = \mathbb{E}\left[W(\varphi)\right]^2 = \langle\varphi, \varphi\rangle_H = \|\varphi\|_H^2.$$

In particular $\int_{[0,1]} \varphi_t \,\mathrm{d}B_t$ is a Gaussian random variable with mean 0 and variance $\|\varphi\|_H^2$.

Finally, notice that span $\left\{\mathbb{1}_{[0,t)}\right\}_{t\geq 0}$ is a dense subspace of $H = L^2([0,1], \mathbb{R})$. Therefore, we can extend the Wiener integral to a linear isometry

$$h \mapsto \int_{[0,1]} h(t) \,\mathrm{d}B_t$$

between $L^2([0,1], \mathbb{R})$ and the Gaussian subspace of $L^2(\Omega)$ spanned by $\left\{W\left(\mathbb{1}_{[0,t)}\right)\right\}_{t\in[0,1]}$. Furthermore, the Wiener integral of $h \in H$ agrees with $W(h)$, and it is a Gaussian random variable with expectation 0 and variance $\|h\|_H^2$.

5.3.2 The Malliavin derivative

In this section, we consider the same Hilbert space $H = L^2([0,1], \mathbb{R})$, and the same probability space as before (Ω, Σ, P), where Σ is the σ-algebra generated by the Brownian motion, $\{B_t\}_{t\geq 0}$, we constructed above. Now denote by $\mathcal{S}$ the set of smooth, cylindrical random variables of the form

$$F = f\left(W(h_1), \cdots, W(h_n)\right),$$

for some $f \in C^\infty(\mathbb{R}^n)$ and $h_i \in H$.

We will now define the Malliavin derivative on $\mathcal{S}$.

Definition 5.3.2. Let $F \in \mathcal{S}$, then we define the *Malliavin derivative*, $\mathscr{D}F$, to be the

H-valued random variable

$$\mathscr{D}_t F := \sum_{1\le i\le n} \frac{\partial f}{\partial x_i}\big(W(h_1),\cdots,W(h_n)\big)\, h_i(t).$$

For example, we have that $\mathscr{D}W(h) = h$, and in particular, we have $\mathscr{D}B_t = \mathbb{1}_{[0,t)}$ for all $t \ge 0$. Furthermore, since S is dense in $L^2(\Omega)$, $\mathscr{D}$ extends uniquely to an operator on all of $L^2(\Omega)$.

On the other hand, if we consider $\mathcal{S}_H$ to be the class of smooth and cylindrical stochastic processes of the form

$$u_t = \sum_{1\le j\le n} F_j h_j(t),$$

for $F_j \in \mathcal{S}$ and $h_j \in H$. Then we can define an analogy to the divergence operator seen previously.

Definition 5.3.3. Let $u \in \mathcal{S}_H$, then we define the *divergence*, δu, of u to be the random variable given by

$$\delta u := \sum_{1\le j\le n}\Big(F_j W(h_j) - \big\langle \mathscr{D}F_j, h_j\big\rangle_H\Big).$$

For example, for $h \in H$, we have that $\delta h = W(h) \in L^2(\Omega)$.

Following our previous discussion of these operators on $\mathbb{R}^n$, it is natural to ask whether they are adjoint, and indeed it turns out that this is the case, as we shall see in the next proposition.

Proposition 5.3.4. *Let $F \in \mathcal{S}$ and $u \in \mathcal{S}_H$, then $\mathscr{D}$ and δ are adjoint. In other words*

$$\mathbb{E}\left(\langle \mathscr{D}F, u\rangle_H\right) = \mathbb{E}\left(F\delta(u)\right).$$

Proof. The proof of this follows a very similar argument to that of Proposition 5.2.1, indeed the statements are almost identical. First we may assume that

$$\begin{aligned} F &= f\left(W(h_1),\cdots,W(h_n)\right),\\ u &= \sum_{1\le j\le n} g_j\left(W(h_1),\cdots,W(h_n)\right) h_j, \end{aligned}$$

where $f, g_j \in C_p^\infty(\mathbb{R}^n)$, and $\{h_j\}_{1\le j\le n}$ are orthonormal elements of H. The orthnormality assumption follows by linearity of W, and the Gram-Schmidt procedure.

So letting $p(x)$ denote the density of the n-dimensional standard Gaussian measure with respect to the Lebesgue measure, we can consider

$$\begin{aligned}
\mathbb{E}\left(\langle \mathscr{D}F, u\rangle_H\right) &= \int_{\mathbb{R}^n} \langle \mathscr{D}F, u\rangle_H \, \mathrm{d}\gamma_n \\
&= \int_{\mathbb{R}^n} \sum_{1\le j,k\le n} g_j(x) \tfrac{\partial f}{\partial x_k}(x) \left\langle h_k, h_j \right\rangle \, \mathrm{d}\gamma_n \\
&= \sum_{1\le j\le n} \int_{\mathbb{R}^n} g_j(x) \tfrac{\partial f}{\partial x_j}(x) p(x) \, \mathrm{d}x \\
&= \sum_{1\le j\le n} \int_{\mathbb{R}^n} f(x) \left[x_j g_j - \tfrac{\partial g_j}{\partial x_j} \right] p(x) \, \mathrm{d}x \\
&= \int_{\mathbb{R}^n} F \delta u \, \mathrm{d}\gamma_n \\
&= \mathbb{E}\left(F\delta(u)\right).
\end{aligned}$$

In the third equality above, we use the orthonormality of $\left\{h_j\right\}_{1\le j\le n}$ to reduce the problem to the n-dimensional case. □

Below, we make note of some basic properties of the derivative and divergence operators.

Proposition 5.3.5. *Let $u, v \in \mathcal{S}_H$, $F \in \mathcal{S}$, and $h \in H$. Then if $\{e_i\}_{i\in\mathbb{N}}$ is an orthonormal basis for H, we have the following*

$$\mathscr{D}_h\left(\delta u\right) = \langle u, h\rangle_H + \delta\left(\mathscr{D}_h u\right); \tag{5.3a}$$

$$\mathbb{E}\left(\delta(u)\delta(v)\right) = \mathbb{E}\left(\langle u, v\rangle_H\right) + \mathbb{E}\left(\sum_{i,j\in\mathbb{N}} \mathscr{D}_{e_i} \left\langle u, e_j\right\rangle_H \mathscr{D}_{e_j} \left\langle v, e_j\right\rangle_H \right). \tag{5.3b}$$

Where we write $\mathscr{D}_h F = \langle \mathscr{D}F, h\rangle_H$.

Proof. To prove (5.3a), suppose that

$$u = \sum_{1\le j\le n} F_j h_j,$$

for some $h_j \in H$. Moreover, noting that $\mathscr{D}_h W(h_j) = \left\langle h_j, h\right\rangle_H$, then we have

$$\begin{aligned}
\mathscr{D}_h\left(\delta u\right) &= \mathscr{D}_h \left(\sum_{1\le j\le n} \left(F_j W(h_j) - \left\langle \mathscr{D}F_j, h_j\right\rangle_H \right) \right) \\
&= \sum_{1\le j\le n} F_j \left\langle h_j, h\right\rangle_H + \sum_{1\le j\le n} \left(\mathscr{D}_h F_j W(h_j) - \left\langle \mathscr{D}_h(\mathscr{D}F_j), h_j\right\rangle_H \right) \\
&= \langle u, h\rangle_H + \delta\left(\mathscr{D}_h u\right).
\end{aligned}$$

To prove (5.3b), first let us rewrite

$$\mathbb{E}\left(\delta(u)\delta(v)\right) = \mathbb{E}\left(\left\langle v, \mathscr{D}(\delta u)\right\rangle_H\right) = \mathbb{E}\left(\sum_{i\in\mathbb{N}} \langle v, e_i\rangle_H \mathscr{D}_{e_i}(\delta u)\right),$$

by Proposition 5.3.4. Then using (5.3a), we have

$$\begin{aligned}
\mathbb{E}\left(\delta(u)\delta(v)\right) &= \mathbb{E}\left(\sum_{i\in\mathbb{N}} \langle v, e_i\rangle_H \left(\langle u, e_i\rangle_H + \delta(\mathscr{D}_{e_i} u)\right)\right) \\
&= \mathbb{E}\left(\langle u, v\rangle_H\right) + \mathbb{E}\left(\sum_{i\in\mathbb{N}} \langle v, e_i\rangle_H \delta(\mathscr{D}_{e_i} u)\right) \\
&= \mathbb{E}\left(\langle u, v\rangle_H\right) + \mathbb{E}\left(\sum_{i,j\in\mathbb{N}} \mathscr{D}_{e_i} \left\langle u, e_j\right\rangle_H \mathscr{D}_{e_j} \left\langle v, e_j\right\rangle_H\right),
\end{aligned}$$

where the last inequality follows from Proposition 5.3.4, similarly to the first step. □

5.4 Wiener chaos

This section aims to construct an orthogonal decomposition of random variables in $L^2(\Omega)$, the so called *Wiener chaos expansion*. We shall do this by using some of the concepts of Malliavin calculus we have developed, in particular, we shall look at the action of the derivative and divergence operators on the Wiener chaos expansion.

5.4.1 Isonormal Gaussian processes and Hermite polynomials

In this section, we shall introduce the basic machinery required to construct an orthogonal decomposition of a probability space. Some aspects have already been introduced previously, but shall be formally restated here, in addition to providing some intuition via examples.

Firstly, we formalise the notion of an isonormal Gaussian process.

Definition 5.4.1. Let H be a real, separable Hilbert space with inner product $\langle\cdot,\cdot\rangle_H$. An *$H$-isonormal Gaussian process* is a family $W = \left\{W(h) : h \in H\right\}$ of real valued random variables defined on a probability space (Ω, Σ, P) such that $W(h)$ is a Gaussian random variable for all $h \in H$ and $\mathbb{E}\left(W(h)W(g)\right) = \langle h, g\rangle_H$ for all $h, g \in H$.

The above is quite a verbose definition, but this is something we have already seen when we defined a map $W : L^2([0,1]) \to L^2(\Omega)$ earlier. In particular, we showed how such a process W could be constructed for an arbitrary separable Hilbert space.

Let us now consider how the Hermite polynomials defined previously interact with isonormal Gaussian processes. First, let us consider the following lemma.

Lemma 5.4.2. *Let X, Y be standard Gaussian random variables, which are jointly Gaussian. Then for $n, m \geq 0$ we have*

$$\mathbb{E}\left(H_n(X)H_m(Y)\right) = \begin{cases} 0 & \text{if } n \neq m \\ n!\mathbb{E}\left(XY\right) & \text{if } n = m \end{cases}.$$

Proof. Notice that

$$\mathbb{E}\left(H_n(X)H_m(Y)\right) = \left\langle H_n(X), H_m(Y)\right\rangle_{L^2(\Omega)}.$$

But, by a similar argument to that used in the proof of Proposition 5.1.4, we have that

$$\left\langle H_n(Y), H_m(Y)\right\rangle_{L^2(\Omega)} = \begin{cases} 0 & \text{if } n \neq m \\ n! \left\langle X, Y\right\rangle_{L^2(\Omega)} & \text{if } n = m \end{cases}.$$

Finally, noting that $\langle X, Y\rangle_{L^2(\Omega)} = \mathbb{E}\left(XY\right)$, completes the proof. □

5.4.2 The Wiener chaos expansion

We are now able to begin constructing an orthogonal decomposition of $L^2(\Omega)$. We shall first define the *Wiener chaos decomposition*, before arguing that it is indeed an orthogonal decomposition of $L^2(\Omega)$. To this end, we define the n-th Wiener chaos as below.

Definition 5.4.3. Let W be an H-isonormal Gaussian process. Then we define the *n-th Wiener chaos*, $\mathscr{H}_n$ to be the closure in $L^2(\Omega, \Sigma, P)$ of the linear span of

$$\left\{H_n\left(W(h)\right) : h \in H,\ \|h\|_H = 1\right\}.$$

In particular, we have that $\mathscr{H}_0$ is the space of constants, since $H_0 = 1$. Moreover, we have that $\mathscr{H}_1 = \left\{W(h) : h \in H\right\} = W$, since $H_1 = x$.

We would now like to argue that this defines an orthogonal decomposition of $L^2(\Omega)$. Before, doing this though, we require the following lemma from [11].

Lemma 5.4.4. *The set $\left\{\exp\left(W(h)\right) : h \in H\right\}$ forms a complete system in $L^2(\Omega, \Sigma, P)$.*

Proof. Take $X \in L^2(\Omega, \Sigma, P)$ such that $\mathbb{E}\left[X \exp\left(W(h)\right)\right] = 0$ for all $h \in H$. Then by linearity of W, we have

$$\mathbb{E}\left[X \exp\left(\sum_{1\leq i\leq n} \alpha_i W(h_i)\right)\right] = 0,$$

for any $\alpha_i \in \mathbb{R}$ and $h_i \in H$. But this implies that the Fourier transform of the measure

$$\nu(A) = \mathbb{E}\left[X\, \mathbb{1}_A\left(W(h_1), \cdots, W(h_n)\right)\right]$$

is 0 on $\mathbb{R}^m$, therefore ν is identically 0. But this implies that for any $B \in \Sigma$, we have $\mathbb{E}\left[X\, \mathbb{1}_B\right] = 0$, implying that $X = 0$. Therefore $\left\{\exp\left(W(h)\right) : h \in H\right\}$ forms a complete system in $L^2(\Omega, \Sigma, P)$. □

Theorem 5.4.5 (Wiener chaos expansion)**.** *Let W be an H-isonormal Gaussian process, then*

$$L^2(\Omega, \Sigma, P) = \bigoplus_{n \in \mathbb{N}_0} \mathscr{H}_n,$$

where Σ is the σ-algebra generated by W.

Proof. Firstly, we would like to show that $\{\mathscr{H}_n\}_{n \in \mathbb{N}_0}$ define orthogonal subspaces. To this end, take $n, m \geq 0$, then since W is a Gaussian process, then for any h, g we have that $W(h), W(g)$ are standard Gaussian random variables, and they are jointly Gaussian. So by Lemma 5.4.2, $H_n\big(W(h)\big)$ and $H_m\big(W(g)\big)$ are orthogonal. Since this holds for arbitrary $h, g \in H$, we have that $\mathscr{H}_n$ and $\mathscr{H}_m$ are orthogonal subspaces of $L^2(\Omega, \Sigma, P)$. Since $n, m \geq 0$ were arbitrary, we conclude that $\{\mathscr{H}_n\}_{n \in N_0}$ consists of orthogonal subspaces of $L^2(\Omega, \Sigma, P)$.

Now it remains to show that for any $F \in L^2(\Omega, \Sigma, P)$, we have that F can be decomposed into elements of $\mathscr{H}_n$. Suppose that this were not the case, in particular, there exists $F \in L^2(\Omega, \Sigma, P)$ such that F is orthogonal to $\bigoplus_{n \in N_0} \mathscr{H}_n$. In particular, since $\mathscr{H}_n$ consists of polynomials of degree n of $W(h)$ for $h \in H$, this implies that F is orthogonal to elements of the form $W(h)^n$ for $n \in N_0$ and $h \in H$. By dominated convergence, this implies that F is orthogonal to $\exp\big(W(h)\big)$. However, $\Big\{\exp\big(W(h)\big)\Big\}_{h \in H}$ is a complete system in $L^2(\Omega, \Sigma, P)$ by Lemma 5.4.4, therefore, $F = 0$, completing the proof. $\square$

It is worth briefly remarking that, denoting by J_n the projection onto the n-th Wiener chaos, Theorem 5.4.5 can be restated for $F \in L^2(\Omega)$ as

$$F = \sum_{n \in N_0} J_n F.$$

5.5 The Ornstein-Uhlenbeck semigroup

Part of the motivation, for our purposes, of the Wiener chaos decomposition is that the derivative and divergence operators defined previously behave nicely, in a sense that we shall discuss, on the n-th Wiener chaos. This will be important as we reintroduce the Ornstein-Uhlenbeck semigroup in a more general setting.

5.5.1 The Ornstein-Uhlenbeck operator

Previously, we defined the Ornstein-Uhlenbeck operator L by (5.2), restated below:

$$Lf := -\delta \mathscr{D} f,$$

where δ and $\mathscr{D}$ were defined on continuous functions from $\mathbb{R}$ into $\mathbb{R}$, $C^1(\mathbb{R})$. We define the Ornstein-Uhlenbeck operator $\mathcal{L}$ as before with definitions of $\mathscr{D}$ and δ given by Definition 5.3.2 and Definition 5.3.3 respectively, in other words for $F \in L^2(\Omega)$

$$\mathcal{L} F = -\delta \mathscr{D} F. \tag{5.4}$$

We change the notation from L to $\mathcal{L}$ in order to differentiate between the two definitions. Specifically, we use L to refer to the definition given by (5.2), and $\mathcal{L}$ to refer to the definition given by (5.4).

Proposition 5.5.1. *Let $F \in L^2(\Omega)$. Then we have*

$$\mathcal{L} J_n F = -n J_n F,$$

where $J_n F$ denotes the projection of F onto $\mathscr{H}_n$.

Before proving the proposition, let us compare this to the action of L on Hermite polynomials. We established that for the n-the Hermite polynomial, we have

$$LH_n = -nH_n.$$

There are obvious parallels with Proposition 5.5.1, where $\mathscr{H}_n$ plays the role of H_n, this seems sensible given our definition of $\mathscr{H}_n$.

Proof. Let $h \in H$, then consider $H_n\left(W(h)\right) \in \mathscr{H}_n$. Then by the properties of Hermite polynomials established previously we have

$$\begin{aligned}
\mathcal{L} H_n\left(W(h)\right) &= -\delta \mathscr{D} H_n\left(W(h)\right) \\
&= -\delta\left(H_n'\left(W(h)\right) h\right) \\
&= -\delta\left(n H_{n-1}\left(W(h)\right) h\right) \\
&= -n H_{n-1}\left(W(h)\right) W(h) + n(n-1) H_{n-2}\left(W(h)\right) \\
&= -n H_n\left(W(h)\right).
\end{aligned}$$

To summarise, we have $\mathcal{L} H_n\left(W(h)\right) = -n H_n\left(W(h)\right)$.

Now we consider, $J_n F$ for $F \in L^2(\Omega)$, then since $J_n F \in \mathscr{H}_n$, the density of the set span$\left\{H_n\left(W(h)\right) : h \in H\right\}$ in $\mathscr{H}_n$, it follows that

$$\mathcal{L} J_n F = -n J_n F.$$

□

The linearity of $\mathcal{L}$ and Theorem 5.4.5 then gives the following corollary.

Corollary 5.5.2. *Let $F \in L^2(\Omega)$, then we can we have the identity*

$$\mathcal{L} F = -\sum_{n \in \mathbb{N}_0} n J_n F.$$

5.5.2 The Ornstein-Uhlenbeck semigroup

We will now formally define the Ornstein-Uhlenbeck semigroup.

Definition 5.5.3. The *Ornstein-Uhlenbeck semigroup* is the one-parameter operator semigroup $\{T_t\}_{t \geq 0}$ defined by

$$T_t F = \sum_{n \in \mathbb{N}_0} e^{-nt} J_n F,$$

for $F \in L^2(\Omega)$.

Now let us consider the following

$$\frac{\mathrm{d}}{\mathrm{d}t} T_t F = \sum_{n\in\mathbb{N}_0} \frac{\mathrm{d}}{\mathrm{d}t} e^{-nt} J_n F = \sum_{n\in\mathbb{N}_0} -ne^{-nt} J_n F = \sum_{n\in\mathbb{N}_0} e^{-nt} \mathcal{L} J_n F = \mathcal{L} T_t F.$$

In other words, we see, as we established for L, that $\mathcal{L}$ is the generator of the Ornstein-Uhlenbeck semigroup.

Let us now provide an equivalent characterisation for the Ornstein-uhlenbeck semigroup. In Proposition 5.1.5 we saw that we could equivalently define the Ornstein-Uhlenbeck semigroup defined on $L^2(\mathbb{R})$ via *Mehler's formula*. This remains true for the Ornstein-Uhlenbeck semigroup as we have defined it on $L^2(\Omega)$, and in fact the proof is almost identical. First recall that $B = \{B_t\}_{t\geq 0}$ is a Brownian motion on (Ω, Σ, P) such that Σ is generated by B.

Proposition 5.5.4 (Mehler's formula)**.** *Let B^1 and B^2 be independent Brownian motions, then for any $F \in L^2(\Omega)$, we have the following identity*

$$T_t F = \mathbb{E}_{B^2}\left(F\left(e^{-t}B^1 + \sqrt{1-e^{-2t}}B^2\right)\right),$$

where $\mathbb{E}_{B^2}$ denotes the expectation with respect to B^2.

Proof. As in Proposition 5.1.5, it suffices to prove the claim for random variables of the form $e^{\lambda W(h)}$, as these form a complete system in $L^2(\Omega)$. To this end, notice that

$$\begin{aligned}
\mathbb{E}_{B^2}\left(\exp\left(e^{-t}B^1 + \sqrt{1-e^{-2t}}B^2\right)\right) &= \exp\left(e^{-t}W(h) + \frac{1}{2}\lambda^2\left(1-e^{-2t}\right)\right) \\
&= e^{\frac{1}{2}\lambda^2} \sum_{n\in\mathbb{N}_0} \frac{e^{-nt}\lambda^n}{n!} H_n\left(W(h)\right) \\
&= T_t e^{\lambda W(h)},
\end{aligned}$$

where we use (5.1) in second and third equalities. $\square$

Conclusion

To conclude this dissertation, let us review what has been achieved. We began by introducing the necessary framework in order to do analysis on Gaussian spaces, in particular, we introduced Gaussian measures, and discussed the structure of a Gaussian space via the Cameron-Martin theorem.

Once we had established the set up, we were then able to present two major results in the form of Fernique's theorem on the exponential tails of Gaussian measures, and the Gaussian isoperimetric inequality. The former provides a strong statement on the integrability of Gaussian measures, and we looked at its application to the infinite dimensional Wiener space. The latter argues that in a Gaussian space, halfspaces maximise measure for a fixed surface measure, which can be compared to the analogous result in Euclidean space, which states that spheres have this property.

The final chapter of this dissertation focused on deriving a calculus for random variables on a Gaussian probability space, by first defining operators on $C(\mathbb{R})$, then by abstracting these definitions to $\mathbb{R}^n$ and finally to the Wiener space. We were then able to explore how these operators acted on certain subspaces, which gave use the Wiener chaos decomposition and gave us a calculus for Gaussian spaces.

In total, this dissertation gives an overview of the Gaussian analysis, with some major theorems presented in a self-contained manner with proofs, in order to provide a resource on the topic.

References

[1] Vladimir Bogachev. *Gaussian Measures*. Vol. 62. American Mathematical Society, Sept. 1998. DOI: 10.1090/surv/062. URL: http://www.ams.org/surv/062.

[2] Vladimir I. Bogachev. "Measure theory". In: *Measure Theory* 1 (2007), pp. 1–575. DOI: 10.1007/978-3-540-34514-5.

[3] Olav Kallenberg. *Foundations of Modern Probability*. eng. 3rd ed. 2021. Vol. 99. Probability Theory and Stochastic Modelling. Cham: Springer International Publishing. ISBN: 3030618706.

[4] R. H. Cameron and W. T. Martin. "Transformations of Wiener integrals under translations". In: *Ann. of Math. (2)* 45 (1944), pp. 386–396. ISSN: 0003-486X. DOI: 10.2307/1969276. URL: https://doi.org/10.2307/1969276.

[5] Xavier Fernique. "Intégrabilité des vecteurs gaussiens". In: *C. R. Acad. Sci. Paris Sér. A-B* 270 (1970), A1698–A1699. ISSN: 0151-0509.

[6] Christer Borell. "The Brunn-Minkowski inequality in Gauss space". In: *Invent. Math.* 30.2 (1975), pp. 207–216. ISSN: 0020-9910. DOI: 10.1007/BF01425510. URL: https://doi.org/10.1007/BF01425510.

[7] V. N. Sudakov and B. S. Cirelson. "Extremal properties of half-spaces for spherically invariant measures". In: *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* 41 (1974). Problems in the theory of probability distributions, II, pp. 14–24, 165.

[8] Antoine Ehrhard. "Symétrisation dans l'espace de Gauss". In: *Math. Scand.* 53.2 (1983), pp. 281–301. ISSN: 0025-5521. DOI: 10.7146/math.scand.a-12035. URL: https://doi.org/10.7146/math.scand.a-12035.

[9] R. Latała. "On some inequalities for Gaussian measures". In: *Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002)*. Higher Ed. Press, Beijing, 2002, pp. 813–822.

[10] Christer Borell. "The Ehrhard inequality". In: *C. R. Math. Acad. Sci. Paris* 337.10 (2003), pp. 663–666. ISSN: 1631-073X. DOI: 10.1016/j.crma.2003.09.031. URL: https://doi.org/10.1016/j.crma.2003.09.031.

[11] David Nualart. *The Malliavin calculus and related topics David Nualart*. eng. Second edition. Probability and its applications. Berlin ; Springer, 2006. ISBN: 3540283293.