Technische Universität München

Department of Mathematics

Bachelor's Thesis

Abstract Wiener Spaces

Gideon Chiusole
from Vienna

Advisor: Prof. Dr. Nina Gantert

M14 - Probability Theory

Submitted on the 30th of August, 2021

I assure the single handed composition of this bachelor's thesis only supported by declared resources.

Gideon Chiusole Garching, 30th of August, 2021

Zusammenfassung

Ziel dieser Arbeit ist es, ein maßtheoretisches und funktionalanalytisches Gerüst eines Differential- und Integralkalküls für unendlichdimensionale, topologische Vektorräume aufzusetzen. Im einführenden Teil betrachten wir ein paar naive Ansätze und stellen schnell fest, dass diese nicht zielführend sind. Anschließend geben wir Motivation aus Teilbereichen der Mathematik und ihren Grenzgebieten, insbesondere aus der konstruktiven Quantenfeldtheorie, dem Malliavinschen Kalkül, der Finanzmathematik & stochastischen partiellen Differentialgleichungen, sowie der Theorie großer Abweichungen.

Im zweiten Teil führen wir die Grundbegriffe von Maßen auf lokalkonvexen topologischen Vektorräumen ein und befassen uns mit dem Problem der Wahl der "richtigen" σ-Algebra. Weiters definieren wir Gaußsche Maße und diskutieren den Satz von Fernique und seine Konsequenzen, z.B. die Einbettung $(E,\tau)^* \hookrightarrow L^p(E,\mu)$. Anschließend definieren und untersuchen wir den zu einem Gaußschen Maß assoziierten Cameron–Martin-Raum und betrachten die Beispiele des $\mathbb{R}^n$ und des klassischen Wienerschen Raumes. Im letzten Abschnitt des Kapitels beweisen wir die Sätze von Cameron und Martin und fassen die Theorie für separable Frechet-Räume zusammen.

Im dritten Teil untersuchen wir den in Abschnitt 1.2 angesprochenen dualen Ansatz, in welchem man von einer formalen Dichte bezüglich eines (hypothetischen) Lebesgue-Maßes ausgeht und den zugehörigen funktionalanalytischen Rahmen konstruiert. Abschließend wenden wir Teile der Theorie auf eine Verallgemeinerung des Satzes von Schilder aus der Theorie der großen Abweichungen an.

Abstract

The goal of this thesis is to set up a measure theoretic and functional analytic framework for a differential and integral calculus on infinite-dimensional topological vector spaces (TVSs). In the introductory part, we make some naive approaches and immediately see why they are doomed to fail. We give further motivation from pure mathematics, physics, and financial economics.

In the second part, we will introduce the basic notions of measures on locally convex TVSs and consider the problem of choosing "the right" sigma-algebra. We then define Gaussian measures and discuss the celebrated Theorem of Fernique and its consequences, e.g. the embedding $(E,\tau)^* \hookrightarrow L^p(E,\mu)$. After that we study the associated Cameron–Martin space and consider the example of finite-dimensional real space $\mathbb{R}^n$ and the classical Wiener space. In the last part of the chapter we state and prove the Theorems of Cameron and Martin and summarize the theory in its most natural setting, separable Frechet spaces.

Finally, we consider the dual viewpoint, mentioned in section 1.2, in which we start from a formal density w.r.t. a (hypothetical) Lebesgue measure and subsequently develop the corresponding functional analytic framework. In the final chapter we employ parts of the theory to obtain a generalized version of the classical Theorem of Schilder from the Theory of Large Deviations.

Contents

1 Introduction

Unless explicitly stated otherwise, all vector spaces are defined over $\mathbb{R}$, locally convex topological vector spaces are assumed to be Hausdorff, and finite-dimensional vectors are column vectors. Frequently used notation and acronyms can be found in appendix B.

1.1 Naive Considerations

The bare minimum for a full-fledged differential and integral calculus is a vector space, a topology, and a sigma-algebra with a measure on it. In the finite-dimensional case, there are canonical choices for each of these: the n-dimensional standard real space, the topology induced by the 2-norm (or any norm, for that matter), and the Borel sigma-algebra with the n-dimensional Lebesgue measure.

Setting	Space	Topology	Measure
Classical	$\mathbb{R}$	induced by $\lvert \cdot \rvert$	$(\mathcal{B}(\mathbb{R}), \lambda)$
Multidimensional	$\mathbb{R}^n$	induced by $\lVert \cdot \rVert$	$(\mathcal{B}(\mathbb{R}^n), \lambda^n)$
Harmonic	$\mathbf{G}$	loc. compact & Hausdorff	$(\mathcal{B}(\mathbf{G}), \mu)$
Inf. dimensional	E	Normed, Frechet, loc. convex, ..., TVS	$(?, ?)$

Up to this point, this arrangement can be seen as a special case of the standard setting of harmonic analysis – a locally compact and Hausdorff topological group, together with the Borel sigma-algebra, and the Haar measure. Under the stated assumptions about the space and the topology, a Haar measure, left or right, always exists and is unique (up to a multiplicative constant). Furthermore, the measure has the crucial (and quasi-defining) property of translation-invariance, i.e.

$$\mu(\cdot) = \mu(\cdot - g), \quad \forall g \in \mathbf{G}.$$

It generalizes the n-dimensional Lebesgue measure. The naive attempt of applying this theory to our setting of infinite-dimensional spaces fails because infinite-dimensional normed spaces are never locally compact, which follows immediately from Riesz's Lemma. We could try using a weaker topology, in the hopes of finding more compact sets and achieving local compactness. This does not work. Not only does not every TVS have a canonical choice for a weaker topology such as the weak* topology (as not every TVS has a pre-dual), but being Hausdorff and locally compact already implies finite-dimensionality for a general TVSs. This leaves the final option of trying to make constructions by hand. However, this fails as well, as the following theorem shows.

Theorem 1.1. *Let $(E, \|\cdot\|)$ be a normed space with $\dim E = \infty$. Then there is no non-trivial, translation-invariant, σ-additive Borel measure μ on $(E, \|\cdot\|)$ s.t. $\mu[B_\varepsilon(0)] < \infty$ for all $\varepsilon > 0$.*

Proof. Since μ is non-trivial, there exists an $A \in \mathcal{B}(E)$ s.t. $\mu(A) > 0$. Hence there exists an $N \in \mathbb{N}$ s.t. $\mu[B_N(0)] > 0$ because otherwise

$$0 < \mu(A) \leq \mu(E) = \mu\left[\bigcup_{n=1}^{\infty} B_n(0)\right] = \lim_{n\to\infty} \mu\big[B_n(0)\big] = 0.$$

Assume w.l.o.g. $N = 1$. Since $\dim E = \infty$, by Riesz's Lemma, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq E$ with $\|x_n\| = 4$ for every $n \in \mathbb{N}$ and $\|x_n - x_m\| \geq 3$ for every $n \neq m$. Thus for every $n \neq m$ we have

$$B_1(x_n) \cap B_1(x_m) = \emptyset \quad \text{and} \quad B_1(x_n) \subseteq B_5(0).$$

Hence by σ-additivity and translation-invariance:

$$\infty = \sum_{n=1}^{\infty} \underbrace{\mu[B_1(x_n)]}_{=\mu[B_1(0)]>0} = \mu\left[\bigcup_{n=1}^{\infty} B_1(x_n)\right] \leq \mu[B_5(0)].$$

A contradiction to $\mu[B_\varepsilon(0)] < \infty$ for every $\varepsilon > 0$. □

1.2 Constructive Quantum Field Theory

See [17, Chap. 20]. In the physics literature on quantum field theory (QFT) one often finds expressions of the form

$$\frac{1}{\beta}\int_E F(\phi)\exp\left\{\frac{i}{\hbar}\mathcal{S}(\phi,0,t)\right\}\mathcal{D}\phi$$

where E is some space of paths or fields, F a functional on E, $\mathcal{S}$ is an action functional, $\beta \in \mathbb{R}$ a normalization constant, and $\mathcal{D}$ is the "infinite-dimensional Lebesgue measure on E". We use quotation marks because, as shown in Theorem 1.1, the latter does not make any sense. The challenge is thus to make rigorous meaning out of the term

$$\frac{1}{\beta}\exp\left\{\frac{i}{\hbar}\mathcal{S}(\phi,0,t)\right\}\mathcal{D}\phi,$$

against which F is integrated, which is formally (!) a measure on a space E defined by a density $\frac{1}{\beta}\exp\left\{\frac{i}{\hbar}\mathcal{S}(\phi,0,t)\right\}$ with respect to $\mathcal{D}$.

Let's consider an explicit example: In his thesis, R. Feynman [11] proposed a formula for the time evolution operator $e^{\frac{-it}{\hbar}\widehat{H}}$ applied to the wave function ψ of a particle as

$$\left(e^{\frac{-it}{\hbar}\widehat{H}}\psi\right)(x_0) = \frac{1}{\beta}\int\limits_{\substack{\text{paths with}\\ x(0)=x_0}} \psi\big(x(t)\big)\exp\left\{\frac{i}{\hbar}\int_0^t\left[\frac{m}{2}|\dot{x}(s)|^2 - V\big(x(s)\big)\right]\mathrm{d}s\right\}\mathcal{D}x \tag{1.1}$$

in which case

$$\mathcal{S}(x,0,t) := \int_0^t\left[\frac{m}{2}|\dot{x}(s)|^2 - V\big(x(s)\big)\right]\mathrm{d}s$$

is the action functional with density $\frac{m}{2}|\dot{x}|^2 - V\big(x(\cdot)\big)$, which is nothing but the Lagrangian $L = T - V$ of the classical trajectories, and $\mathcal{D}$ the "infinite-dimensional Lebesgue measure" on a space of paths. Applying a Wick rotation $t \mapsto -it$, i.e. passing from a problem in Minkowski space to one in Euclidean space, makes the exponent real and allows us to rewrite (1.1) as

$$\int\limits_{\substack{\text{paths with}\\ x(0)=x_0}} \psi\big(x(t)\big) \exp\left\{ -\frac{1}{\hbar}\int_0^t V\big(x(s)\big)\mathrm{d}s \right\} \underbrace{\frac{1}{\beta}\exp\left\{ -\frac{1}{\hbar}\int_0^t \frac{m}{2}|\dot{x}(s)|^2 \mathrm{d}s \right\}\mathcal{D}x}_{\mathrm{d}\mu(x)}.$$

Now, the bilinear form in the kinetic energy part makes the latter factor look like a (real-valued) Gaussian density, suggesting that it is at least plausible that we can rigorously interpret μ as a (Gaussian) probability measure on $C_{x_0}([0,t],\mathbb{R}^n)$ and subsequently recover the original problem by analytic continuation. Indeed, one can show the following theorem.

Theorem 1.2. *(Feynman–Kac Formula, [17, Thm. 20.3]) Let $V : \mathbb{R}^3 \to \mathbb{R}$ s.t. $V = f + b$ with $f \in L^2(\mathbb{R}^3)$ and $b \in L^\infty(\mathbb{R}^3)$. Then for any $x_0 \in \mathbb{R}^3$*

$$\left(e^{\frac{-t}{\hbar}\widehat{H}}\psi\right)(x_0) = \int\limits_{C_{x_0}([0,t],\mathbb{R}^3)} \psi\big(x(t)\big)\exp\left\{ -\frac{1}{\hbar}\int_0^t V\big(x(s)\big)\mathrm{d}s\right\}\mathrm{d}\mu^\sigma_{x_0}(x)$$

where $\mu^\sigma_{x_0}$ is the Wiener measure on $C_{x_0}([0,t],\mathbb{R}^3)$ with variance $\sigma := \frac{\hbar}{m}$.

Another example is ϕ^4 theory, where (the Minkowski version of) a path integral of the form

$$\frac{1}{\beta}\int_{\mathcal{F}^n} F(\phi)\exp\left\{ -\frac{1}{\hbar}\int_{\mathbb{R}^n} \beta_1\|\nabla\phi(x)\|^2 + \beta_2\phi(x)^2 + \beta_3\phi(x)^4 dx\right\}\mathcal{D}\phi$$

is considered. Here, $\beta_1, \beta_2, \beta_3 \in \mathbb{R}$ are constants and F is some functional on a space of fields $\mathcal{F}^n$. Many more examples can be found in [25].

Takeaway Many interesting measures are given as formal expression on some space of (generalized) functions; which space might not be clear a priori. The Feynman–Kac formula solves the prototypical case and suggests Gaussian measures as the rigorous interpretation of those expressions.

1.3 Malliavin Calculus

See [13]. Let $(C_0[0,1], H_0^1[0,1], \mu)$ be the classical Wiener space as defined in section 4.1 and let

$$W(\tilde{h}) : C_0[0,1] \to \mathbb{R},\ \omega \mapsto \int_0^1 h(s)\mathrm{d}\omega(s)$$

denote the Paley–Wiener integral[1] for $h \in L^2[0,1]$ and $\tilde{h}(t) := \int_0^t h(s)\mathrm{d}s \in H_0^1[0,1]$. We would like to define a differential calculus for functionals of Brownian motion, i.e. for measurable functions $C_0[0,1] \to \mathbb{R}$ which are defined only μ-a.s. So, for example,

$$S_t = S_0 \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right\} = S_0 \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma \int_0^1 1_{[0,t]}(s)\mathrm{d}B_s\right\}$$

at a fixed time $t > 0$.[2] Such a differential operator should extend the action

$$W\big(\tilde{h}\big) \mapsto \tilde{h}, \quad \tilde{h} \in H_0^1[0,1]$$

to more general functionals $F : C_0[0,1] \to \mathbb{R}$. In a sense, the operator should differentiate F w.r.t. B_t. Akin to the weak derivative on an open interval $(a,b) \subseteq \mathbb{R}$, we would like to do so by defining such an operator on a space D_0 of well-behaved functions, yielding a closable operator, and then extending it to a closed operator, defined on a larger domain $D_{\max}$. In the case of the former, the derivative operator $\phi \mapsto \phi'$ is defined on $D_0 := C_c^\infty(a,b)$, mapping to $L^2(a,b)$, and then extended to $D_{\max} = H^1(0,1)$, giving the weak derivative defined on the first Hilbert–Sobolev space. For the Malliavin derivative we choose

$$D_0 := \left\{F := p\Big(W(\widetilde{h_1}), \ldots, W(\widetilde{h_n})\Big) : p \in \mathbb{R}[x_1, \ldots, x_n], \widetilde{h_1}, \ldots, \widetilde{h_n} \in H_0^1[0,1], n \in \mathbb{N}\right\}$$

i.e. the space of polynomial functions evaluated after Paley–Wiener integrals. The definition of the Frechet derivative (which applies to well-behaved and in particular everywhere defined functionals) suggests that the right notion of directional derivative $\partial_\varphi F : C_0[0,1] \to \mathbb{R}$ of F in the direction of $\varphi \in C_0[0,1]$ is

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} F\big(\omega + \varepsilon\varphi\big) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} \frac{|F(\omega) - F(\omega + \varepsilon\varphi)|}{\varepsilon}, \quad \omega \in C_0[0,1]. \tag{1.2}$$

evaluated at $\varepsilon = 0$. However, recall that $W(\widetilde{h_i})$ is defined as a μ-a.s. limit. Hence $W(\widetilde{h_i})$ and thus F is only well-defined on a set $\mathrm{Dom}(F) \subseteq C_0[0,1]$ with $\mu\big[\mathrm{Dom}(F)\big] = 1$. However, by the classical Cameron–Martin Theorem (which will be presented in a more general form in section 3.4), unless φ lies precisely in $H_0^1[0,1]$,

$$\mu\Big[\mathrm{Dom}\Big(F(\cdot + \varepsilon\varphi)\Big)\Big] = \mu\Big[\mathrm{Dom}(F) - \varepsilon\varphi\Big] = \mu_{\varepsilon\varphi}\Big[\mathrm{Dom}(F)\Big] = 0.$$

This makes (1.2) μ-a.s. ill-defined and thus not suited for our purpose. A better approach is to only consider derivatives in Cameron–Martin directions:

[1] The Paley–Wiener integral as we define it here is sometimes denoted $W(h)$, where $h \in L^2[0,1]$ as defined above. The difference is merely notational, since $h \mapsto \tilde{h}$ is a linear isometric isomorphism from $L^2[0,1]$ to $H_0^1[0,1]$. In subsection 3.3.3 we will realize W as a map $H_0^1[0,1] \to L^2(C_0[0,1])$, i.e. from the Cameron–Martin space into the space of square integrable functionals on $C_0[0,1]$, which justifies the notation used here.

[2] It is the (unique and strong) solution to the SDE

$$dS_t = S_t\big(\mu \ \mathrm{d}t + \sigma \ \mathrm{d}B_t\big)$$

and models the price of an underlying in the Black–Scholes model at a fixed time $t > 0$ where $\mu \in \mathbb{R}$, $\sigma > 0$. See section 1.4

$$\partial_{\widetilde{h}} F(\omega) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} F(\omega + \varepsilon \widetilde{h}), \quad \widetilde{h} \in H_0^1[0,1].$$

Restricting to $h_i := 1_{[0,t_i]}$ in the definition of F and applying the chain rule twice gives

$$\sum_{i=1}^{n} \partial_i p\Big(W\big(\widetilde{1_{[0,t_1]}}\big), \dots, W\big(\widetilde{1_{[0,t_n]}}\big)\Big) \int_0^{t_i} h(s)\mathrm{d}s = \langle DF, h\rangle_{L^2[0,1]} = \langle \widetilde{DF}, \widetilde{h}\rangle_{H_0^1[0,1]},$$

where we define

$$DF = \sum_{i=1}^{n} \partial_i p\Big(W\big(\widetilde{1_{[0,t_1]}}\big), \dots, W\big(\widetilde{1_{[0,t_n]}}\big)\Big) 1_{[0,t_i]}, \quad t_1, \dots, t_n \in [0,1].$$

This then naturally extends to

$$DF = \sum_{i=1}^{n} \partial_i p\Big(W\big(\widetilde{h_1}\big), \dots, W\big(\widetilde{h_n}\big)\Big) h_i, \quad h_1, \dots, h_n \in L^2[0,1],$$

which finally gives DF as a linear operator $C_0[0,1] \to L^2[0,1]$. Comparing this to the derivative operator in finite dimensions, it suggests that $L^2[0,1]$ (or equivalently $H_0^1[0,1]$) should take the role of a ("stochastic") tangent space to points in $C_0[0,1]$.

Takeaway Constructions that rely on a Gaussian measure are deeply connected to the Cameron–Martin space. While Gaussian measures on finite-dimensional spaces are quasi-invariant w.r.t. translation in any direction (i.e. translation yields a possibly different, but equivalent measure) on infinite-dimensional spaces, they have a strong tendency to become mutually singular, which has to be taken into account.

1.4 Mathematical Finance & Stochastic PDE

See [24] and [26]. Let S_t denote the price of a stock S at time $t \geq 0$. If the price at t_0 is known, then the price at a later time $t_0 + T$ can be approximated by

$$S_{t_0+T} = S_{t_0} + T\mu S_{t_0},$$

where μ is the expected return of S, which can be rewritten as

$$\frac{S_{t_0+T} - S_{t_0}}{T} = \mu S_{t_0}.$$

Letting $T \to 0$, this leads to a simple model of the stock price via the ODE

$$\frac{dS_t}{\mathrm{d}t} = \mu S_t. \tag{1.3}$$

However, plotting solutions to this equation yields figure 2, which does not coincide with what we can observe on the stock market. Firstly, this model suggests that $(S_t)_{t\geq 0}$ is deterministic, i.e. that the evolution of the price can be predicted with certainty. This is not the case. Secondly, it implies that $(S_t)_{t\geq 0}$ is smooth, which is also not accurate. The reason for these two deficiencies of the model is that we have assumed μ to be **the** return

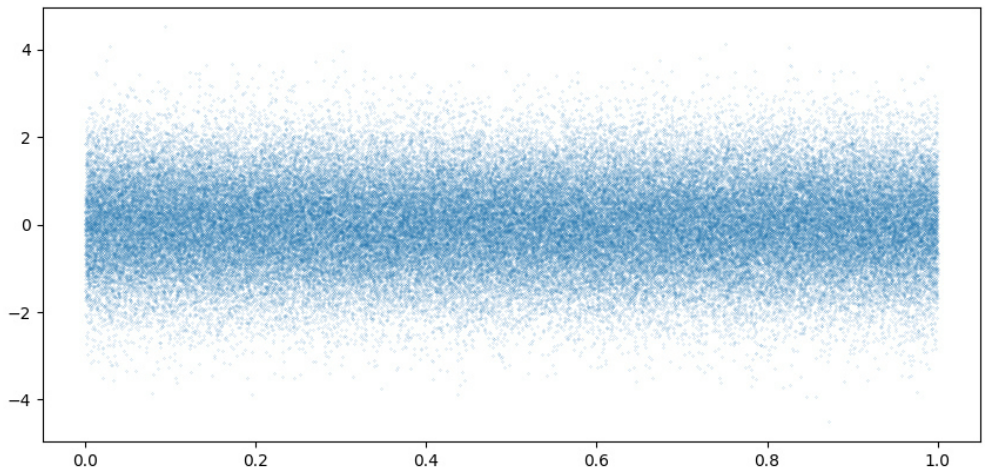

Figure 1: Sample of white noise.

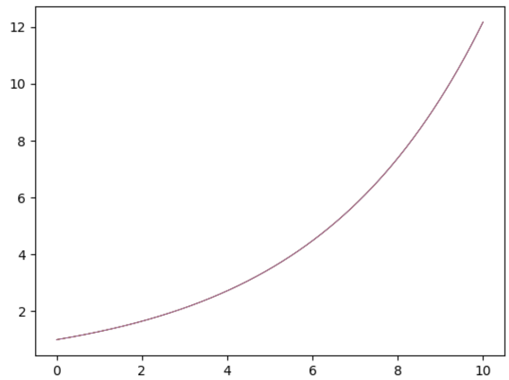

Figure 2: Solution to (1.3).

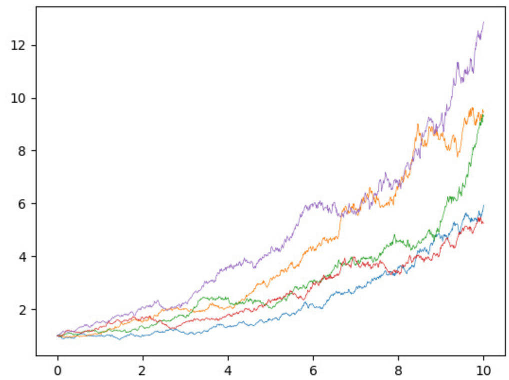

Figure 3: Solution to (1.4).

of S, which stays perfectly constant over time and is deterministic. A more accurate assumption is that μ is close to constant, but nonetheless continuously influenced by some noise $(\eta_t)_{t\geq 0}$ in the market. Empirically measuring these disturbances, we observe that for any $s, t \geq 0$ s.t. $s \neq t$ the random variables η_s and η_t are independent and have distribution $\mathcal{N}(0,1)$. Such a family of random variables is called **white noise**.
Introducing a parameter $\sigma > 0$ which controls the volatility of the stock S yields $\sigma\eta_t \sim \mathcal{N}(0, \sigma^2)$ and the ODE becomes

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = (\mu + \sigma\eta_t)S_t. \tag{1.4}$$

Plotting solutions to the above equation leads to figure 3. The graph looks very rough[3]

[3]Indeed, solutions to (1.4), called **geometric Brownian motion**, are almost surely nowhere differentiable.

and suggests that it cannot satisfy a differential equation in the classical sense.[4] One way to make sense out of (1.4) is the Itô interpretation. It asserts that a stochastic process $(S_t)_{t\geq 0}$ is a solution to (1.4) if it is adapted and satisfies

$$S_t = S_0 + \int_0^t \mu S_s \mathrm{d}s + \int_0^t \sigma S_s \mathrm{d}B_s$$

where, heuristically speaking, $(B_t)_{t\geq 0}$ is some stochastic process whose derivative is η_t, and, precisely speaking, $(B_t)_{t\geq 0}$ is standard Brownian motion. More generally, stochastic ODEs have the form

$$\mathrm{d}X_t = b(t, X_t)\ \mathrm{d}t + \sigma(t, X_t)\ \mathrm{d}\underline{B}_t, \quad X_0 = \xi,$$

where $b : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$, $(\underline{B}_t)_{t\geq 0}$ is d-dimensional Brownian motion, ξ is an n-dimensional random vector which is independent of $(\underline{B}_t)_{t\geq 0}$, and $(X_t)_{t\geq 0}$ is a (stochastic) process to be solved for. In this more general context, $(\underline{B}_t)_{t\geq 0}$ should be a stochastic process "whose derivative is white noise with values in whatever space s.t. $\sigma(t, X_t)\mathrm{d}\underline{B}_t$ takes values in the space in which $X_t(\omega)$ lies"; in this case the white noise should take values in $\mathbb{R}^d$ s.t. $\sigma(t, X_t)\mathrm{d}\underline{B}_t$ lies in $\mathbb{R}^n$, which is where $X_t(\omega)$ lies. This situation can still be treated with the conventional tools from standard stochastic analysis. Namely, define d-dimensional Brownian motion as $\underline{B}_t = (B_t^{(1)}, \dots, B_t^{(d)}) \in \mathbb{R}^d$ for every $t \geq 0$ where $(B_t^{(1)})_{t\geq 0}, \dots, (B_t^{(d)})_{t\geq 0}$ are independent 1-dimensional Brownian motions. However, now consider the case of a stochastic PDE, for example

$$\mathrm{d}X_t = (\Delta - m^2)X_t\ \mathrm{d}t + \mathrm{d}B_t.$$

Here, the space in which $X_t(\omega)$ lies is a function space, and thus $(B_t)_{t\geq 0}$ should be Hilbert/Banach space valued, which needs to be rigorously defined and cannot be treated with the classical tools.

Takeaway For the purposes of ODEs and SDEs, Brownian motion is a way to deal with white noise. If the white noise takes value in $\mathbb{R}$ or $\mathbb{R}^n$, then the classical theory is sufficient. With a view towards white noise with values in more general topological vector spaces, it is natural to study Brownian motion abstractly.

1.5 Large Deviations

See [7] and see chapter 5 for brief summary of the main idea of Large Deviations.
Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of $\mathbb{R}^n$-valued iid random variables with $X_1 \sim \mathcal{N}(0, 1^{n\times n})$. Then by Cramer's Theorem, the sequence $(\mu_N)_{N\in\mathbb{N}}$, defined by

$$\mu_N := \mathcal{L}\left(\frac{1}{N}\sum_{n=1}^N X_n\right), \quad \forall N \in \mathbb{N},$$

satisfies a large deviation principle with good rate function

[4] In fact, the process $(\eta_t)_{t\geq 0}$ we hypothesized before does not exist in the sense that a generic realization $\big(\eta_t(\omega)\big)_{t\geq 0}$ is not a measurable function [19, Exmp. 1.2.5.] Hence the right hand side of equation (1.4) is not measurable, and thus $(S_t)_{t\geq 0}$ cannot be differentiable in the classical sense.

$$I(x) = \frac{1}{2}\|x\|^2_{\mathbb{R}^d}, \quad x \in \mathbb{R}^n.$$

This suggests that the Large Deviations are controlled by the Euclidean norm of $x \in \mathbb{R}^n$. However, the true nature becomes apparent when we consider a general covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$ that may differ from the identity matrix $1^{d\times d}$. Consider a sequence as above but with $X_1 \sim \mathcal{N}(0,\Sigma)$. Then $(\mu_N)_{N\in\mathbb{N}}$ still satisfies a large deviation principle, but I takes the form

$$I(x) = \frac{1}{2}\langle x, \Sigma^{-1}x\rangle_{\mathbb{R}^n}, \quad x \in \mathbb{R}^n.$$

What we notice is that it is not the Euclidean norm that is important, but the norm weighted by Σ^{-1}, which is the Cameron–Martin norm. A similar result is true for Gaussian measure on infinite-dimensional spaces (see chapter 5). In fact, there we will see that I is finite only on the Cameron–Martin space itself, not on the entire space.

Takeaway Properties of Gaussian measures largely depend only on the properties of their Cameron–Martin space, and not on the entire space they are defined on. In the words of L. Gross:

> "However, it only became apparent with the work of I. E. Segal, dealing with the normal distribution on a real Hilbert space, that the role of [the Cameron–Martin space] was indeed central, and that in so far as analysis on [the entire space] is concerned, the role of [the entire space] itself was auxiliary for many of Cameron and Martin's Theorems, and in some instances even unnecessary." [15]

2 Measures on Locally Convex Spaces

2.1 Choice of Sigma Algebra

As opposed to the finite-dimensional situation, it is not clear what the "correct" sigma-algebra on E should be.[5] Various concepts of the theory of Gaussian measures, such as their very definition, are built upon continuous linear functionals. In particular, any sensible choice of sigma-algebra on (E,τ) should at least make those measurable. In other words, the sigma-algebra $\mathcal{A}$ on E should be chosen (as weak as possible) with the property that $(E,\tau)^* \hookrightarrow L^0(E,\mathcal{A})$ is well-defined.

Definition 2.1. Let (E,τ) be a locally convex TVS, $A \in \mathcal{B}(\mathbb{R}^n)$ a Borel set, and $f_1,\dots,f_n \in (E,\tau)^*$. Then

$$C_{A,f_1,\dots,f_n} := \Big\{x \in E : \big(f_1(x),\dots,f_n(x)\big) \in A\Big\} = (f_1,\dots,f_n)^{-1}(A)$$

is called a **cylinder set**, A is called its **basis**, and $f_1,\dots,f_n$ its **generators**.

[5] The truth is rather that on $\mathbb{R}^n$ all of the reasonable choices simply coincide.

Proposition 2.2. *Let (E, τ) be a locally convex TVS. Then the collection of all cylinder sets of (E, τ), i.e.*

$$\mathcal{C}(E) := \Big\{ C_{A,f_1,\dots,f_n} : n \in \mathbb{N}, A \in \mathcal{B}(\mathbb{R}^n), f_1, \dots, f_n \in E^* \Big\}$$

forms an algebra of sets (although not necessarily a sigma-algebra).

Proof. We show $\{\complement, \cap\}$-stability:
Let $A, B \in \mathcal{B}(\mathbb{R}^n)$ and $f_1, \dots, f_n, g_1, \dots, g_m \in E^*$ be arbitrary. Then

$$C_{A,f_1,\dots,f_n}^{\complement} = \Big\{ x \in E : \big(f_1(x), \dots, f_n(x)\big) \in A^{\complement} \Big\} = C_{A^{\complement},f_1,\dots,f_n} \in \mathcal{C}(E)$$

and

$$\begin{aligned} C_{A,f_1,\dots,f_n} \cap C_{B,g_1,\dots,g_m} &= \Big\{ x \in E : \big(f_1(x), \dots, f_n(x)\big) \in A, \big(g_1(x), \dots, g_m(x)\big) \in B \Big\} \\ &= \Big\{ x \in E : \big(f_1(x), \dots, f_n(x), g_1(x), \dots, g_m(x)\big) \in A \times B \Big\} \\ &= C_{A\times B,f_1,\dots,f_n,g_1,\dots,g_m} \in \mathcal{C}(E). \end{aligned}$$

□

Proposition 2.3. *Let (E, τ) be a locally convex TVS. Then the sigma-algebra generated by $\mathcal{C}(E)$ coincides with the smallest sigma-algebra making all $f \in E^*$ measurable, i.e. $\sigma(\mathcal{C}(E)) = \sigma(\mathcal{C}_0(E))$ where*

$$\mathcal{C}_0(E) := \Big\{ f^{-1}(A) : A \in \mathcal{B}(\mathbb{R}), f \in E^* \Big\}.$$

$\sigma(\mathcal{C}(E))$ will thus also be called the ***weak sigma-algebra****.*

Proof. We have $\sigma(\mathcal{C}_0(E)) \subseteq \sigma(\mathcal{C}(E))$ by definition. To show the other inclusion, note that $\sigma(\mathcal{C}(E))$ is the smallest sigma-algebra making all the functionals $(f_1, \dots, f_n)$ with $f_1, \dots, f_n \in E^*$ measurable. Hence it is sufficient to show that $\sigma(\mathcal{C}_0(E))$ makes those measurable as well. So let $F = (f_1, \dots, f_n)$ be as above. Then by the universal property of the product sigma-algebra on $\mathbb{R}^n$ (which coincides with $\mathcal{B}(\mathbb{R}^n)$) the function F is measurable if and only if for every $1 \le i \le n$ the function f_i is measurable. This is the case by assumption of $\sigma(\mathcal{C}_0(E))$. Hence $\sigma(\mathcal{C}_0(E))$ is a sigma-algebra making all functionals of the form $(f_1, \dots, f_n)$ measurable, and thus contains the smallest sigma-algebra with this property, which is $\sigma(\mathcal{C}(E))$. □

Remark 2.4. Note that, as opposed to $\mathcal{C}(E)$, the family $\mathcal{C}_0(E)$ is not $\cap$-stable. In the proof of $\cap$-stability of Proposition 2.2 the sets A and B lie in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, but $A \times B \subseteq \mathbb{R}^{n+m}$. Hence $\mathcal{C}(E)$ is more suitable for Dynkin-type arguments (see for example the proof of Theorem 2.5).

Back to the problem of choice of sigma-algebra. We have already established that the continuous linear functionals should be measurable. Is that sufficient?

On the one hand, one would like to be able to make arguments as in Theorem 2.15, by testing properties of μ against bounded linear functionals on E. This suggests the sigma-algebra $\sigma(\mathcal{C}(E))$. On the other hand, with this choice, there may be continuous (non-linear) functionals which are not measurable. To make those measurable as well we would choose the Baire sigma-algebra $\sigma\big(E; C(E,\tau)\big)$. And yet, neither of the two may coincide with the Borel sigma-algebra $\sigma(\tau)$, which is arguably the natural choice all along. To summarize, we have the following inclusions:

$$\sigma(\mathcal{C}(E)) = \sigma(E; E^*) \subseteq \sigma\Big(E; C(E,\tau)\Big) \subseteq \sigma(\tau) \tag{2.1}$$

denoting in ascending order

- $\sigma(\mathcal{C}(E))$ the smallest sigma-algebra containing $\mathcal{C}(E)$ (which coincides with $\sigma(E; E^*)$ by Proposition 2.3),
- $\sigma\big(E; C(E,\tau)\big)$ the smallest sigma-algebra making all continuous (possibly non-linear) functionals on (E,τ) measurable,
- $\sigma(\tau)$ the Borel sigma-algebra on (E,τ).

Fortunately though, not only for finite-dimensional TVS, but for separable Frechet spaces all four coincide and even more is true:

Theorem 2.5. *Let (E,τ) be separable Frechet. Then for any family $F \subseteq E^*$ that separates points we have*

$$\sigma(E; F) = \sigma(\mathcal{C}(E)) = \sigma\Big(E; C(E,\tau)\Big) = \sigma(\tau) \tag{2.2}$$

Moreover, there exists a countable sub-family $F_0 \subseteq F$ which also separates points, i.e. for any $x \neq y \in E$ there is a $f_0 \in F$ s.t. $f_0(x) \neq f_0(y)$, and for which $\sigma(E; F_0)$ coincides with (2.2).

Proof. We want to show $\sigma(\tau) \subseteq \sigma(E; F_0)$. Then the nesting in (2.1) and $F_0 \subseteq F \subseteq E^*$, and hence $\sigma(E; F_0) \subseteq \sigma(E; F) \subseteq \sigma(\mathcal{C}_0(E))$, imply the result. A family $F \subseteq E^*$ always exists since (E,τ) is assumed to be Frechet;[6] e.g. take $F = E^*$.

(1) Firstly we will show that there exists a countable subfamily $F_0 \subseteq F$ that also separates points. For every $f \in F$ define

$$U_f := \Big\{(x,y) \in E \times E : f(x) \neq f(y)\Big\} \subseteq E \times E,$$

which is the set of points $x, y \in E$ which are separated by f. As the complement of the pre-image of $\{0\}$ under the continuous function $(x,y) \mapsto f(x) - f(y)$, this set is open. Since (E,τ) is separable and metrizable, so is $(E \times E, \tau \otimes \tau)$, which is thus second countable. Let $\{V_n\}_{n\in\mathbb{N}}$ denote a countable basis of $(E \times E, \tau \otimes \tau)$. Then

$$\bigcup_{f\in F} U_f = \bigcup_{f\in F} \cup_{n=1}^{\infty} V_{f,n} = \bigcup_{k\in\mathbb{N}} V_k = \bigcup_{f_k\in F_0} U_{f_k}. \tag{2.3}$$

For the second equality, choose a re-indexing of $F \times \mathbb{N}$ by $\mathbb{N}$ - this is possible since we have a countable basis. In the third equality, define F_0 by choosing for each $k \in \mathbb{N}$ a

[6]See [31, Kor. VIII.2.13.].

functional $f_k \in E^*$ s.t. $V_k \subseteq U_{f_k}$ - this is possible because, by construction, every V_k is a subset of U_f for some $f \in E^*$. Formula (2.3) tells us that the set of points separated by F (which is all of E) agrees with the set being separated by F_0. Thus F_0 also separates points.

(2) Let $\{B_r(q) : q \in \mathbb{Q}, r \in \mathbb{Q}^+\}$ be the usual countable basis of the standard topology on $\mathbb{R}$. Then

$$\mathcal{E} := \Big\{ f^{-1}\big(B_r(q)\big) : q \in \mathbb{Q}, r \in \mathbb{Q}^+, f \in F_0 \Big\}$$

separate points in E, i.e. for any $x \neq y \in E$ there is a $A \in \mathcal{E}$ s.t. $x \in A$ and $y \notin A$. To see this, let $x \neq y \in E$ and let $f \in F_0$ s.t. $f(x) \neq f(y)$. Then choose $q \in \mathbb{Q}, r \in \mathbb{Q}^+$ s.t.

$$\big|q - f(x)\big| < \frac{1}{4}\big|f(x) - f(y)\big| \quad \text{and} \quad \frac{1}{4}\big|f(x) - f(y)\big| < r < \frac{1}{2}\big|f(x) - f(y)\big|.$$

Then $x \in f^{-1}\big(B_r(q)\big)$ but $y \notin f^{-1}\big(B_r(q)\big)$.

(3) Choose an indexing of $\mathcal{E}$ by $\mathbb{N}$ and define the function

$$g(x) = \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}_{f^{-1}\left(B_r(q)\right)}(x), \quad x \in E$$

as a point-wise limit. Since the sets of $\mathcal{E}$ separate points, the function is injective (if $g(x) = g(y)$ then x and y must lie in the exact same sets of the form $f^{-1}\big(B_r(q)\big)$ - a contradiction to separation). Since the $f \in F_0$ are continuous and g is a point-wise limit of indicator functions on pre-images of f, the function is $\sigma(\tau) - \mathcal{B}(\mathbb{R})$-measurable and also $\sigma(E; F_0) - \mathcal{B}(\mathbb{R})$-measurable.

(4) Let $B \in \sigma(\tau)$ be arbitrary. Then we show $B \in \sigma(E; F)$. By a Theorem of Lusin–Suslin [21, Chap. 15, Thm. 15.1], since g is Borel measurable between Polish spaces and injective, $g(B) \in \mathcal{B}(\mathbb{R})$. Note that this is not true for a general measurable g; the assumption of injectivity and thus of $\mathcal{E}$ separating points is needed here. Thus, since g is injective, B equals $g^{-1}\big(g(B)\big)$, which lies in $\sigma(E; F_0)$ because g is $\sigma(E; F_0) - \mathcal{B}(\mathbb{R})$-measurable. We conclude $\sigma(\tau) \subseteq \sigma(E; F_0)$, which completes the proof. □

The above theorem is not true for general locally convex TVSs. Consider the following example:

Example 2.6. Let $\mathbb{R}^{[0,1]}$ be the set of real-valued functions on the unit interval $[0,1]$ equipped with the locally convex topology τ induced by the functionals $\{\mathrm{ev}_t : t \in [0,1]\}$, i.e. with the usual product topology. On the one hand, since τ is Hausdorff, $\{0\}$ is closed and thus Borel-measurable. Hence $\{0\} \in \sigma(\tau)$. On the other hand, $\sigma(\mathcal{C}(\mathbb{R}^{[0,1]}))$ is contained in the sigma-algebra consisting of sets of the form

$$\Xi_{k \in \mathbb{N}} \Big(\Xi_{j \in \mathbb{N}} \big(\dots \Xi_{i \in \mathbb{N}} A_i \big)_j \Big)_k$$

where $\Xi \in \{\cup, \cap\}$ and $A_i \in \mathcal{C}(\mathbb{R}^{[0,1]})$. But

$$\{0\} = \bigcap_{t\in[0,1]} \mathrm{ev}_t^{-1}\left(\{0\}\right),$$

which is not of the form above, since it is an uncountable intersection. Hence $\sigma(\tau) \neq \sigma(\mathcal{C}(\mathbb{R}^{[0,1]}))$. While $\mathbb{R}^{[0,1]}$ is separable, it is not Frechet. This can be seen for example by noting that $\mathbb{R}^{[0,1]}$ is not first countable.

2.2 Momenta

Akin to the finite-dimensional theory, the (first few) moments of a measure carry a lot of information about it. We will see that for Gaussian measures, the first two momenta are enough to characterize it. A naive definition of the mean $\mathfrak{m}_\mu$ of a measure μ would be an element of E s.t.

$$\mathfrak{m}_\mu = \int_X x \mathrm{d}\mu(x).$$

However, as opposed to measures on $\mathbb{R}$, this integral is not in the sense of Lebesgue, but in a (generalized) sense of Bochner - a crucial difference. Instead, we make the following definition, which also generalizes the mean of a measure on $\mathbb{R}^n$.

Definition 2.7. Let (E,τ) be a locally convex TVS and μ a measure on $\sigma(E;E^*)$. Then define the **mean** $\mathfrak{m}_\mu : E^* \to \mathbb{R}$ via

$$\mathfrak{m}_\mu(f) := \int_E f(x)\mathrm{d}\mu(x), \quad \forall f \in E^*$$

and the **covariance form** $\mathfrak{q}_\mu : E^* \times E^* \to \mathbb{R}$ via

$$\mathfrak{q}_\mu(f,g) := \int_E \Big[f(x) - \mathfrak{m}_\mu(f)\Big]\Big[g(x) - \mathfrak{m}_\mu(g)\Big]\mathrm{d}\mu(x), \quad \forall f,g \in E^*.$$

If it is clear from the context we will just write $\mathfrak{m}$ and $\mathfrak{q}$. If $\mathfrak{m}_\mu = 0$, then μ is called **centred**. If there exists a non-zero $f \in E^*$ s.t. $\mathfrak{q}(f,f) = 0$, then μ is called **degenerate**. If $\mathfrak{q}$ is non-degenerate, then it induces a norm $\|\cdot\|_{\mathfrak{q}}$ defined by $\sqrt{\mathfrak{q}(\cdot,\cdot)}$.

Remark 2.8. If every $f \in E^*$ is 1-integrable, then $\mathfrak{m}$ is well-defined and an element of the algebraic dual space $(E^*)'$ of E^*, which assigns to every f its expected valued in the usual sense, i.e. $\mathfrak{m}(f) = \mathbb{E}(f)$. However, it is a priori not clear in which sense and whether at all $\mathfrak{m}$ is continuous and when $\mathfrak{m}$ is representable as an evaluation functional ev_a for some $a \in E$ - in general (assuming the continuum hypothesis) neither is true (see [2, Thm. 2.12.2.]). A failure of $\mathfrak{m}_\mu$ corresponding to an element in E is somewhat pathological - think of the Strong Law of Large Numbers of Ranga Rao 5.2 - and will be avoided from section 3.3 onwards.

Remark 2.9. The covariance form $\mathfrak{q}_\mu$, if it exists, is positive semi-definite, symmetric and bilinear; and as such it is uniquely determined by its values $\mathfrak{q}(f,f)$ for any $f \in E^*$, since for any $f,g \in E^*$

$$\mathfrak{q}(f,g) = \frac{1}{2}\big(\mathfrak{q}(f+g,f+g) - \mathfrak{q}(f,f) - \mathfrak{q}(g,g)\big).$$

Since $\mathfrak{q}$ is a bilinear form, it is representable by a linear operator $\mathfrak{C}_\mu : E^* \to (E^*)'$ via

$$\forall f, g \in E^* : \quad [\mathfrak{C}_\mu(f)](g) = \mathfrak{q}(f, g).$$

The operator $\mathfrak{C}_\mu$ is referred to as the **covariance operator** of μ. When there is no risk of confusion we will just write $\mathfrak{C}$. Similar to the case of $\mathfrak{m}$ it is natural to ask whether $\mathfrak{q}$ is continuous and when $\mathfrak{C}$ is given by evaluations, i.e. when $\mathfrak{C} : E^* \to E \subseteq (E^*)'$. The continuity and representability of $\mathfrak{m}$ and $\mathfrak{C}$ will be discussed in subsection 3.2.3.

Finite-dimensional case: For $E = \mathbb{R}^n$, we have $(\mathbb{R}^n)^* \simeq \mathbb{R}^n$ via

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \cdot \right\rangle \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

and also $\big((\mathbb{R}^n)^*\big)' \simeq \mathbb{R}^{n**} \simeq \mathbb{R}^n$ via

$$\left\langle \cdot, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\rangle \mapsto \left\langle \cdot, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\rangle \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Hence the mean of a measure μ on $\mathbb{R}^n$ is an element $\mathfrak{m} \in \mathbb{R}^n$ s.t. for every $y \in \mathbb{R}^n$

$$\langle y, \mathfrak{m} \rangle = \int_{\mathbb{R}^n} \langle y, x \rangle \mathrm{d}\mu(x) = \sum_{i=1}^n y_i \int_{\mathbb{R}^n} x_i \mathrm{d}\mu(x) = \left\langle y, \int_{\mathbb{R}^n} x \mathrm{d}\mu(x) \right\rangle.$$

Hence

$$\mathfrak{m} = \begin{pmatrix} \int_{\mathbb{R}^n} x_1 \mathrm{d}\mu(x) \\ \vdots \\ \int_{\mathbb{R}^n} x_n \mathrm{d}\mu(x) \end{pmatrix}.$$

For the covariance form we have for any $x, y \in \mathbb{R}^n$

$$\begin{aligned} \mathfrak{q}(\langle x, \cdot \rangle, \langle y, \cdot \rangle) &= \int_{\mathbb{R}^n} \Big[\langle x, z \rangle - \mathfrak{m}\big(\langle x, \cdot \rangle\big) \Big] \Big[\langle y, z \rangle - \mathfrak{m}\big(\langle y, \cdot \rangle\big) \Big] \mathrm{d}\mu(z) \\ &= \sum_{i,j=1}^n \int_{\mathbb{R}^n} \Big[x_i z_i - x_i \mathfrak{m}_i \Big] \Big[y_j z_j - y_j \mathfrak{m}_j \Big] \mathrm{d}\mu(z) \\ &= \sum_{i,j=1}^n x_i y_j \int_{\mathbb{R}^n} \Big[z_i - \mathfrak{m}_i \Big] \Big[z_j - \mathfrak{m}_j \Big] \mathrm{d}\mu(z) \\ &= \langle x, \Sigma y \rangle, \end{aligned}$$

where Σ is the usual covariance matrix of the Gaussian measure μ defined by

$$\Sigma_{i,j} = \int_{\mathbb{R}^n} \big[z_i - \mathfrak{m}_i \big] \big[z_j - \mathfrak{m}_j \big] \mathrm{d}\mu(z), \quad 1 \le i, j \le n,$$

or, if we interpret μ as the distribution of a random variable $Z : \Omega \to \mathbb{R}^n$,

$$\Sigma_{i,j} = \mathbb{E}\big[Z_i - \mathbb{E}(Z_i)\big]\big[Z_j - \mathbb{E}(Z_j)\big], \quad 1 \le i, j \le d.$$

We see that in the finite-dimensional case $\mathfrak{m}$ and $\mathfrak{q}$ are continuous and both are representable as evaluations in the sense described above. The covariance operator $\mathfrak{C}$ coincides with the usual covariance matrix Σ after the canonical identification of $\mathbb{R}^n$ with its dual.

The above situation can be generalized to Banach spaces.

Proposition 2.10. *Let $(E, \|\cdot\|_E)$ be a Banach space and μ a measure on $\sigma(E; E^*)$. If the mean exists in the sense of Bochner, i.e. if $x \mapsto x$ is integrable, then $\mathfrak{m}_\mu$ lies in the continuous dual of E^* and is representable by $a := \int_E x \mathrm{d}\mu(x)$.*

Proof. To see this, note that linear operators can be pulled inside Bochner integrals. Hence for any $f \in E^*$ we have

$$\mathrm{ev}_a(f) = f(a) = f\left(\int_E x \, \mathrm{d}\mu(x)\right) = \int_E f(x) \, \mathrm{d}\mu(x) = \mathfrak{m}(f).$$

Since E^* separates points, $a \in E$ is unique with this property. □

2.3 Characteristic Functional

Harmonic analysis in the sense of Pontryagin takes place on locally compact Hausdorff groups. We are not in this setting. However, under some assumptions, some theorems can be recovered. In particular, finite measures on $\sigma(E; E^*)$ and also Radon measures on $\sigma(\tau)$ are characterized by their Fourier transform, justifying the name of characteristic functional.

Definition 2.11. Let (E, τ) be a locally convex TVS and μ a finite signed measure on $\sigma(E; E^*)$. Then $\widehat{\mu} : E^* \to \mathbb{C}$, defined by

$$\widehat{\mu}(f) := \int_E \exp\big[if(x)\big]\mathrm{d}\mu(x), \quad f \in E^*,$$

is called the **characteristic functional** or **Fourier transform** of μ. For $E = \mathbb{R}^n$ this coincides with the usual Fourier transform of μ in the distributional sense. The Fourier transform describes a linear operator from the space $\mathbf{M}\big(E, \sigma(E; E^*)\big)$ of finite signed measures into the space of complex-valued functionals on E^*.

The integral in the above definition is well-defined since the integrand is bounded by 1 and μ is assumed to be finite. Note also that $\widehat{\mu}$ is non-linear except for trivial cases.

Proposition 2.12. *Let (E, τ) be a locally convex TVS, μ a finite signed measure on $\sigma(E; E^*)$, and $\widehat{\mu}$ its characteristic functional. Then the following are true:*

1. *$|\widehat{\mu}(f)| \le \widehat{|\mu|}(0) = |\mu|(E)$ for any $f \in E^*$. Hence $\widehat{\mu}$ is bounded, and bounded by 1 if μ is a probability measure.*
2. *$\overline{\widehat{\mu}(f)} = \hat{\mu}(-f)$ for any $f \in E^*$.*

3. *If μ is tight, then $\widehat{\mu}$ is continuous w.r.t. the topology of compact convergence.*
4. *If (E, τ) is a separable Banach space, then $\widehat{\mu}$ is continuous w.r.t. the operator norm topology.*
5. *If $\mu \geq 0$, then $\widehat{\mu}$ is positive definite, i.e. for any $f_1, \ldots, f_n \in E^*, a_1, \ldots, a_n \in \mathbb{C}$:*

$$\sum_{i,j=1}^{n} \widehat{\mu}(f_i - f_j) a_i \overline{a_j} \geq 0.$$

Proof. 1. Let $f \in E^*$ be arbitrary. Then

$$|\widehat{\mu}(f)| = \left| \int_E \exp(if(x)) \mathrm{d}\mu(x) \right| \leq \int_E \underbrace{\left| \exp\big(if(x)\big) \right|}_{=1} \mathrm{d}|\mu|(x) = \widehat{|\mu|}(0) = |\mu|(E)$$

2. Clear.
3. Let $f_0 \in E^*$ and $\varepsilon > 0$ be arbitrary. We want to find a τ-compact set $K \subseteq E$ and $\delta > 0$ s.t. $B(K, \delta, f_0) \subseteq \widehat{\mu}^{-1}\Big(B_\varepsilon\big(\widehat{\mu}(f_0)\big)\Big)$ where

$$B(K, \delta, f_0) := \Big\{ f \in E^* : \sup_{x \in K} \big| f(x) - f_0(x) \big| < \delta \Big\}.$$

By tightness of μ choose $K \subseteq E$ s.t. $|\mu|\big(E \setminus K\big) \leq \varepsilon/4$. By uniform continuity of $t \mapsto \exp(it)$ choose a $\delta > 0$ s.t.

$$|\exp(it) - \exp(is)| \leq \frac{\varepsilon}{2|\mu|(K)}, \quad \text{when} \quad |t - s| < \delta.$$

Then for any $f \in E^*$ with

$$\sup_{x \in K} \big| f(x) - f_0(x) \big| < \delta$$

i.e. for any $f \in B(K, \delta, f_0)$ we have

$$\begin{aligned} |\widehat{\mu}(f) - \widehat{\mu}(f_0)| &\leq \int_E \big| \exp\big\{if(x)\big\} - \exp\big\{if_0(x)\big\} \big| \mathrm{d}|\mu|(x) \\ &\leq 2 \cdot |\mu|\big(E \setminus K\big) + \int_K \big| \exp\big\{if(x)\big\} - \exp\big\{if_0(x)\big\} \big| \mathrm{d}|\mu|(x) \\ &\leq 2 \cdot |\mu|\big(E \setminus K\big) + \varepsilon/2 \leq \varepsilon. \end{aligned}$$

i.e. $\widehat{\mu}(f) \in B_\varepsilon\big(\widehat{\mu}(f_0)\big)$, which shows the claim.

Note that the characteristic functional may not be weak*-continuous. One can show sequential weak*-continuity using the Dominated Convergence Theorem, but that does not imply continuity, e.g. weak* duals of reflexive Banach spaces are never sequential.

4. Since E is separable Banach, $\sigma(E;E^*) = \sigma(\tau)$ by Theorem 2.5. Thus, since E is Polish, every finite signed measure μ on $\sigma(E;E^*)$ is tight by Ulam's Tightness Theorem. Since on E^* the topology of compact convergence is weaker than the topology induced by the operator norm (i.e. the norm of bounded convergence) the result follows from (3).
5. Let $\mu \geq 0$ and $f_1, \ldots, f_n \in E^*, a_1, \ldots, a_n \in \mathbb{C}$ be arbitrary. Then

$$\begin{aligned}\sum_{i,j=1}^{n} \widehat{\mu}(f_i - f_j) a_i \overline{a_j} &= \int_E \sum_{i,j=1}^{n} \exp\Big\{ i\big(f_i(x) - f_j(x)\big) \Big\} a_i \overline{a_j} \, \mathrm{d}\mu(x) \\ &= \int_E \sum_{i,j=1}^{n} \exp\big\{ i f_i(x) \big\} a_i \, \overline{\exp\big\{ i f_j(x) \big\} a_j} \, \mathrm{d}\mu(x) \\ &= \int_E \underbrace{\left| \sum_{j=1}^{n} \exp\big\{ i f_j(x) \big\} a_j \right|^2}_{\geq 0} \mathrm{d}\mu(x) \geq 0\end{aligned}$$

□

Recall Bochner's Theorem for the Fourier transform of probability measures on $(\mathbb{R}^n, \|\cdot\|)$ (or any locally compact abelian Hausdorff group):

Theorem 2.13. *(Bochner) Let $\phi : \mathbb{R}^n \to \mathbb{C}$ be a positive-definite and continuous function s.t. $\phi(0) = 1$. Then there exists a unique probability measure μ on $\mathbb{R}^n$ s.t. $\phi = \widehat{\mu}$.*

Proof. See [27, Thm. IX.9] for the case of $\mathbb{R}^n$ and [12, Thm. 4.19] for the general case of a locally compact abelian Hausdorff group. □

For E a separable Hilbert space there exists a generalization.

Theorem 2.14. *(Minlos–Sazanov) Let $(E, \langle\cdot,\cdot\rangle)$ be a separable Hilbert space and $\phi : E \to \mathbb{C}$ a positive definite complex-valued continuous functional. Then the following are equivalent:*

- *ϕ is the characteristic functional of a finite Borel measure on E.*
- *There is a symmetric trace class linear operator S on E s.t. ϕ is continuous w.r.t. the norm $\|x\| := \sqrt{\langle x, Sx\rangle}$ for every $x \in E$.*

Proof. See [18, Chap. I, Thm. 4.4]. □

Although there are some positive results (e.g. [23]), there is no suitable analogue for Bochner's Theorem when E is not a Hilbert space; not even when E is Banach.

In general, the Fourier transform only characterizes a measure on $\sigma(E;E^*)$, not on $\sigma(\tau)$. Under some mild assumptions, however, this can be resolved.

Theorem 2.15. *(Uniqueness of the Fourier transform) Let (E,τ) be a locally convex TVS and μ_1 and μ_2 are finite signed measures on $\sigma(E;E^*)$ s.t. $\widehat{\mu_1} = \widehat{\mu_2}$.*

(1) Then $\mu_1 = \mu_2$.

(2) If μ_1 and μ_2 are defined on $\sigma(\tau)$ and Radon, then they agree not only on $\sigma(E; E^*)$, but all of $\sigma(\tau)$.

(3) If (E, τ) is separable Frechet, then μ_1 and μ_2 are defined on, and agree on $\sigma(\tau) = \sigma(E; E^*)$.

Proof. (1) We want to show that for any $\mu \in \mathbf{M}\big(E, \sigma(E; E^*)\big)$ we have that $\widehat{\mu} = 0$ implies $\mu = 0$. Since the Fourier transform is a linear operator on $\mathbf{M}\big(E, \sigma(E; E^*)\big)$, this implies the result. Let $f_1, \ldots, f_n \in E^*$ be arbitrary and define $F : E \to \mathbb{R}^n$ via $F(x) = (f_1(x), \ldots, f_n(x))$, $\forall x \in E$. Denote by $\phi_F : E^* \to \mathbb{C}$ the characteristic functional of F as a random variable. Then for every $\xi \in \mathbb{R}^n$ we have

$$\begin{aligned}
\phi_F(\xi) &= \int_{\mathbb{R}^n} \exp\Big\{ i\langle \xi, y\rangle \Big\} \, \mathrm{d}\Big[\mu \circ F^{-1}\Big](y) \\
&= \int_E \exp\Big\{ i\langle \xi, F(x)\rangle \Big\} \, \mathrm{d}\mu(x) \\
&= \int_E \exp\left\{ i \sum_{i=1}^n \xi_i f_i(x) \right\} \, \mathrm{d}\mu(x) \\
&= \prod_{i=1}^n \int_E \exp\Big\{ i\xi_i f_i(x) \Big\} \, \mathrm{d}\mu(x) \\
&= \prod_{i=1}^n \widehat{\mu}(\xi_i f_i) = 0
\end{aligned}$$

where the latter term is 0 by assumption. Hence in particular the law of the random vector F is the 0-measure, i.e. for every $A \in B(\mathbb{R}^n)$ we have

$$0 = \Big[\mu \circ F^{-1}\Big](A) = \mu(F \in A) \tag{2.4}$$

Since the $f_1, \ldots, f_n \in E^*$ were arbitrary, this implies that μ agrees with the zero measure on every cylinder set $F^{-1}(A)$. Since the algebra of cylinder sets is in particular $\cap$-stable, the Dynkin's Theorem implies that μ is 0 on the sigma-algebra generated by the cylinder sets $\sigma(\mathcal{C}(E))$, which proves the claim.

(2) Let E be locally convex. Then E is Hausdorff and completely regular [29, Chap. 1, 1.3], and E^* separates points of E. Thus if μ_1 and μ_2 are assumed to be defined on $\sigma(\tau)$ and Radon, then $\mu_1 = \mu_2$ on $\sigma(\tau)$ by [30, Chap. IV, Thm. 2.2 (b)].

(3) If E is separable Frechet, $\sigma(E; E^*)$ agrees with $\sigma(\tau)$ by Theorem 2.5, and the claim follows from (1). $\square$

3 Gaussian Measures

3.1 First Definitions

Definition 3.1. Let (E, τ) be a locally convex TVS. A measure μ on $\sigma(E; E^*)$ is called **Gaussian** if for all $f \in E^*$ the push-forward measure $\mu \circ f^{-1}$ is Gaussian on $\mathbb{R}$, i.e. if it has density

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R},$$

for some $a \in \mathbb{R}, \sigma \geq 0$. In other words, we require all $f \in E^*$ to be Gaussian random variables in the usual sense.

In this definition we consider Dirac measures as Gaussian, i.e. we allow for a Gaussian random variable to be constant a.s. Such distributions will be called **degenerate**. A Gaussian measure is called **non-degenerate** if all $f \in E^*$ are non-degenerate. This is equivalent to $\mathfrak{q}$ being non-degenerate as defined in Definition 2.7 since

$$\mathfrak{q}(f,f) = \int_E \Big[f(x) - \mathbb{E}(f)\Big]^2 \mathrm{d}\mu(x) = \int_{\mathbb{R}} \Big[y - \mathbb{E}(f)\Big]^2 \mathrm{d} \underbrace{\Big[\mu \circ f^{-1}\Big]}_{=\mathcal{N}\left(\mathbb{E}(f), \mathrm{Var}(f)\right)} (y) = \mathrm{Var}(f).$$

which is 0 if and only if f is constant and equal to $\mathbb{E}(f)$ a.s.

Remark 3.2. It immediately follows from the definition that for any $f_1, \dots, f_n \in E^*$ the vector $(f_1, \dots, f_n)$ is a Gaussian random vector in $\mathbb{R}^n$ for any $n \in \mathbb{N}$. To see this, let $\alpha \in \mathbb{R}^n$. Then $\langle \alpha, (f_1, \dots, f_n)\rangle = \sum_{i=1}^n \alpha_i f_i$ lies in E^* and is thus Gaussian by assumption. Hence $(f_1, \dots, f_n)$ is a Gaussian random vector.

Many important Gaussian measures arise naturally as the distribution of function space valued random variables (random processes). We will give some examples in subsection 4.4.

Example 3.3. Let $(\mathbb{R}^{\mathbb{N}}, \tau)$ be the set of real-valued sequences equipped with the product topology. This is a Frechet space and its dual can naturally be identified with the space $c_{00}(\mathbb{R})$ of real-valued, eventually 0 sequences via the pairing

$$\langle (a_n)_{n\in\mathbb{N}}, (x_n)_{n\in\mathbb{N}}\rangle = \sum_{n=1}^{\infty} a_n x_n, \quad (a_n)_{n\in\mathbb{N}} \in c_{00}(\mathbb{R}), (x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}.$$

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real-valued, iid, $\mathcal{N}(0,1)$-distributed random variables defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Since for any $(a_n)_{n\in\mathbb{N}} \in c_{00}(\mathbb{R})$ the function

$$\Omega \to \mathbb{R}^{\mathbb{N}} \to \mathbb{R}, \quad \text{defined by} \quad \omega \mapsto (X_1(\omega), X_2(\omega), \dots) \mapsto \sum_{n=1}^{\infty} a_n X_n(\omega)$$

is measurable, so is $\Omega \to \mathbb{R}^{\mathbb{N}}$ w.r.t the sigma-algebra on $\mathbb{R}^{\mathbb{N}}$ generated by $c_{00}(\mathbb{R})$. Thus we may interpret the discrete time stochastic process $(X_n)_{n\in\mathbb{N}}$ as a random variable with values in $(\mathbb{R}^{\mathbb{N}}, \tau)$. Now define μ on $\sigma(\mathbb{R}^{\mathbb{N}}; c_{00}(\mathbb{R}))$ as the distribution of this process. μ is a probability measure since $(X_n)_{n\in\mathbb{N}}$ takes valued in $\mathbb{R}^{\mathbb{N}}$ $\mathbb{P}$-a.s. and a Gaussian measure since, as a finite sum of independent Gaussian random variables, $\sum_{n=1}^{\infty} a_n X_n$ is Gaussian as well for any $(a_n)_{n\in\mathbb{N}} \in c_{00}(\mathbb{R})$.

Note that for a Gaussian measure μ, the operator $\mathfrak{m}_\mu$ and the form $\mathfrak{q}_\mu$ are always well-defined. In fact, $\mathfrak{m}_\mu$ and $\mathfrak{q}_\mu$ characterize μ completely via the Fourier transform.

Theorem 3.4 (Characterization of Gaussian measures by their characteristic functional)**.** *Let (E, τ) be a locally convex TVS and μ a probability measure on $\sigma(E; E^*)$. Then μ is Gaussian if and only if there exists a positive, semi-definite, bilinear form $\mathfrak{q} : E^* \times E^* \to \mathbb{R}$ and a linear form $\mathfrak{m} : E^* \to \mathbb{R}$ s.t.*

$$\widehat{\mu}(f) = \exp\left\{ i\mathfrak{m}(f) - \frac{1}{2}\mathfrak{q}(f, f) \right\}, \quad \forall f \in E^*.$$

Proof. Let μ be a probability measure on $\sigma(E; E^*)$.
"$\Rightarrow$" Assume μ is Gaussian. Then by definition, every $f \in E^*$ is a Gaussian random variable $E \to \mathbb{R}$. Hence

$$\widehat{\mu}(f) = \int_E \exp\{if(x)\} \mathrm{d}\mu(x) = \int_{\mathbb{R}} e^{iy} \underbrace{\mathrm{d}\left[\mu \circ f^{-1}\right](y)}_{=\mathrm{d}\mathcal{N}\left(\mathbb{E}(f), \mathrm{Var}(f)\right)(y)} = \phi_f(1) = \exp\left\{ i\mathbb{E}(f) - \frac{1}{2}\,\mathrm{Var}(f) \right\}$$

where ϕ_f denotes the characteristic function of f as a random variable. So $\mathfrak{q}$ and $\mathfrak{m}$ are the required forms by Definition 2.7 and Remark 2.9.
"$\Leftarrow$" Let $f \in E^*$ be arbitrary. Then, for every $t \in \mathbb{R}$

$$\phi_f(t) = \int_E \exp\{itf(x)\} \mathrm{d}\mu(x) = \exp\left\{ i\mathfrak{m}(tf) - \frac{1}{2}\mathfrak{q}(tf, tf) \right\} = \exp\left\{ it\mathfrak{m}(f) - \frac{1}{2}t^2\mathfrak{q}(f, f) \right\}$$

i.e. the characteristic function of f coincides with that of a $\mathcal{N}\big(\mathbb{E}(f), \mathrm{Var}(f)\big)$-distributed random variable. Thus, by Bochner's Theorem, f is Gaussian, and thus μ is Gaussian. $\square$

Note that this does not mean that for every $\mathfrak{q}$ and $\mathfrak{m}$ as above there exists a Gaussian measure on E having $\exp\left\{ i\mathfrak{m}(f) - \frac{1}{2}\mathfrak{q}(f, f) \right\}$ as its characteristic functional. The theorem only makes a statement about the type of an already existing measure on E, not about its existence. For example, Proposition 3.10 shows that if E is a Banach space $\mathfrak{q}$ needs to be representable by a compact linear operator $\mathfrak{C} : E^* \to E$. So for instance, on an infinite-dimensional Hilbert space $\big(H, \langle \cdot, \cdot \rangle_H\big)$ the functional $\exp\left\{ -\frac{1}{2}\langle f, f \rangle_H \right\}$ is not the characteristic functional of any countably additive measure (see also [2, Cor. 2.3.2.]). We will consider this example in more depth in Proposition 4.6 and Proposition 4.9.

3.2 Fernique's Theorem

3.2.1 Tail Estimates for Gaussian Measures

On finite-dimensional spaces, Gaussian measures have very good properties w.r.t. integration due to their exponential decay. This translates well to the infinite-dimensional case and is the subject of the following celebrated Theorem of Fernique.

Theorem 3.5 (Fernique [10]). *Let (E, τ) be a locally convex TVS and μ a centred Gaussian measure on $\sigma(E; E^*)$. Let $S : E \to [0, \infty)$ be a semi-norm that is $\sigma(E; E^*) - \mathcal{B}(\mathbb{R})$-measurable.*

(1) Then there exists an $\alpha > 0$ s.t.

$$\int_E \exp\left\{\alpha S(x)^2\right\} \mathrm{d}\mu(x) < \infty \tag{3.1}$$

Sharp bounds can be found in [2, Thm. 2.8.5.].

(2) In particular, for any $\lambda > 0$

$$\int_E \exp\left\{\lambda S(x)\right\} \mathrm{d}\mu(x) < \infty.$$

(3) If $(E, \|\cdot\|_E)$ is a Banach space, then there exist constants $\varepsilon, C > 0$ s.t. for all $t \geq 0$

$$\mu\Big(\big\{x \in E : \|x\|_E \geq t\big\}\Big) \leq Ce^{-\varepsilon t^2} \tag{3.2}$$

In other words, the norm has Gaussian tails. In particular, for every $\varepsilon' < \varepsilon$

$$\int_E \exp\left\{\varepsilon' \|x\|_E^2\right\} \mathrm{d}\mu(x) < \infty$$

and for any $p > 0$

$$\int_E \|x\|_E^p \mathrm{d}\mu(x) < \infty. \tag{3.3}$$

Proof. (1) See [2, Thm. 2.8.5.].

(2) Let $\alpha > 0$ satisfy equation (3.1). Then for any $\lambda > 0$

$$\begin{aligned}\int_E e^{\lambda S(x)} \mathrm{d}\mu(x) &\leq \int_{\left\{\lambda > \alpha S(x)\right\}} e^{\lambda S(x)}\, \mathrm{d}\mu(x) + \int_{\left\{\lambda \leq \alpha S(x)\right\}} e^{\lambda S(x)}\, \mathrm{d}\mu(x) \\ &\leq e^{\lambda \frac{\lambda}{\alpha}} + \int_E e^{\alpha S(x)^2}\, \mathrm{d}\mu(x) < \infty.\end{aligned}$$

(3) For a direct proof and tail bounds see [8, Thm. 4.10]. The theorem still holds under surprisingly mild assumptions on the measure μ (see [16, Prop. 3.10, Thm. 3.11]) For the first corollary, note that

$$\begin{aligned}\int_E \exp\left\{\varepsilon'\|x\|_E^2\right\}\mathrm{d}\mu(x) &= \int_0^\infty \mu\Big\{x\in E : \exp\left\{\varepsilon'\|x\|_E^2\right\} > t\Big\}\mathrm{d}t \\ &= \int_0^\infty \mu\left\{x\in E : \|x\|_E^2 > \frac{\ln t}{\varepsilon'}\right\}\mathrm{d}t \\ &= 1 + \int_1^\infty \mu\left\{x\in E : \|x\|_E > \sqrt{\frac{\ln t}{\varepsilon'}}\right\}\mathrm{d}t \\ &\le 1 + \int_1^\infty t^{-\frac{\varepsilon}{\varepsilon'}}\mathrm{d}t < \infty,\end{aligned}$$

where the ε in the last line is the one from equation (3.2). By choice of ε' the exponent is strictly smaller than -1, which makes the integral finite.

To see formula (3.3), note that $t^p \in O(\exp\{\varepsilon' t^2\})$, i.e. there exists a $t_0 \ge 0, B > 0$ s.t. for every $t \ge t_0$ the inequality $t^p \le \exp\{\varepsilon' t^2\}$ holds. □

3.2.2 Embeddings of $(E,\tau)^* \hookrightarrow L^p(E,\mu)$

The prime consequence of Fernique's Theorem is the fact that for a Gaussian measure μ on a locally convex TVS E the bounded linear functionals can be seen as elements of $L^p(E,\mu)$ with $1 \le p < \infty$.

Proposition 3.6. *Let (E,τ) be a locally convex TVS, μ a centred Gaussian measure on $\sigma(E;E^*)$, and $1 \le p < \infty$.*

(1) Then the inclusion $\jmath : (E,\tau)^* \to (L^p,\mu)$ is linear and well-defined. If μ is non-degenerate, then $\jmath$ is injective.

(2) If E is Banach, then $\jmath$ is a compact linear operator with operator norm $\|\jmath\| \le \left(\int_E \|x\|_E^p \,\mathrm{d}\mu(x)\right)^{\frac{1}{p}}$.

Proof. (1) Linearity is clear. Well-definedness follows from Theorem 3.5. If μ is non-degenerate, then $\mathfrak{q}$ is a non-degenerate bilinear form, and thus, if $\jmath(f) = 0$, then

$$0 = \|\jmath(f)\|_p^p = \int_E |f(x)|^p \mathrm{d}\mu(x),$$

and thus $f = 0$ μ-a.s. Hence

$$\mathfrak{q}(f,f) = \int_E |f(x)|^2 \mathrm{d}\mu(x) = 0 \tag{3.4}$$

and thus $f = 0$.

(2) Linearity is clear. For the boundedness let $f \in E^*$ be arbitrary. Then

$$\|\jmath(f)\|_p^p = \int_E |f(x)|^p \mathrm{d}\mu(x) \le \|f\|_{E^*}^p \int_E \|x\|_E^p \mathrm{d}\mu(x) \tag{3.5}$$

and taking the p-th root and the supremum over all $f \in E^*$ gives the result. The operator norm of $\mathtt{j}$ is the L^p norm of $\|\cdot\|_E : E \to \mathbb{R}$.

In order to see compactness, let $(f_n)_{n\in\mathbb{N}}$ be a bounded sequence in E^*. Assume w.l.o.g. $\forall n \in \mathbb{N} : \|f_n\|_{E^*} \leq 1$. Then by Banach–Alaoglu there exists a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ converging weak* (i.e. point-wise) to some $f \in E^*$ with $\|f\|_{E^*} \leq 1$. Hence by the Dominated Convergence Theorem, justified by

$$|f_{n_k}(x)| \leq \underbrace{\|f_{n_k}\|_{E^*}}_{\leq 1} \|x\|_E \leq \|x\|_E$$

and Fernique's Theorem, the limit also exists in an L^p-sense:

$$\lim_{n\to\infty} \int_E |f_{n_k}(x) - f(x)|^p \mathrm{d}\mu(x) = 0$$

□

For $p = 2$, where the norm is induced by the $\mathfrak{q}$-form, $\mathtt{j}$ being bounded implies that the $\mathfrak{q}$-norm on E^* is weaker than the operator norm. Better yet, $\mathtt{j}$ being a compact operator implies that the $\mathfrak{q}$-norm is "a lot" weaker than the operator norm.

Remark 3.7. I am not entirely sure about the continuity of $\mathtt{j}$ when E is just locally convex. The question is whether $f \mapsto \int_E |f(x)|^p \mathrm{d}\mu(x)$ is continuous, i.e. if the p-th moment is continuous. For $p = 1, 2$ Theorem 3.8 and 3.9 give a criterion for this, and since Gaussian measures have exponential decay, I would expect those criteria to be true for $p \geq 3$ as well.

3.2.3 Representability of $\mathfrak{m}$ and $\mathfrak{q}$

Another consequence of Fernique's Theorem are the following two criteria for when $\mathfrak{m}$ and $\mathfrak{q}$ are representable.

Theorem 3.8. *Let (E, τ) a locally convex TVS and μ a Gaussian measure on $\sigma(E; E^*)$.*

(1) The mean $\mathfrak{m} : (E, \tau)^* \to \mathbb{R}$ is representable by an evaluation functional $\mathfrak{m} = ev_a$ for some $a \in E$ if and only if it is continuous w.r.t. the weak topology on $(E, \tau)^*$.*

(2) For any $g \in (E, \tau)^*$ the functional $\mathfrak{C}(g) : (E, \tau)^* \to \mathbb{R}$ is representable by an evaluation functional if and only if it is weak-continuous; in which case $\mathfrak{C}$ takes values in $E \subseteq E^{**}$.*

(3) Both assumptions are satisfied if (E, τ) is a separable Banach space.

Proof. (1) & (2) Since (E, τ) is locally convex there exists a canonical bijection between E and the continuous dual of E^* equipped with the weak* topology via $x \mapsto \mathrm{ev}_x$ (see [4, Chap. V, Thm. 1.3]).

(3) To show that the assumption is satisfied if E is a separable Banach space, note firstly that it is enough to show sequential continuity by [4, Chap. V, Cor. 12.8]. Now note that weak* convergence implies

$$\sup_{n\in\mathbb{N}} |f_n(x)| < \infty, \quad \forall x \in E,$$

and thus $\sup_{n\in\mathbb{N}} \|f_n\|_{E^*} =: \alpha < \infty$ by the Uniform Boundedness Principle. By Fernique's Theorem

$$\int_E \big|\sup_{n\in\mathbb{N}} f_n(x)\big| \mathrm{d}\mu(x) \le \int_E \sup_{n\in\mathbb{N}} \big|f_n(x)\big| \mathrm{d}\mu(x) \le \int_E \sup_{n\in\mathbb{N}} \|f_n\|_{E^*} \|x\|_E \,\mathrm{d}\mu(x) \le \alpha \underbrace{\int_E \|x\|_{E^*}}_{<\infty} < \infty,$$

i.e. the sequence $(f_n)_{n\in\mathbb{N}}$ is dominated by an integrable function $\sup_{n\in\mathbb{N}} f_n(x)$. Hence the Dominated Convergence Theorem gives

$$\mathfrak{m}(f_n) = \int_E f_n(x)\mathrm{d}\mu(x) \to \int_E f(x)\mathrm{d}\mu(x) = \mathfrak{m}(f).$$

The case of $\mathfrak{C}$ follows similarly. □

The proof rests on the Weak Representation Theorem $(E^*, \text{weak}^*)^* \simeq E$ (see [4, Chap. V, Thm. 1.3]). The Mackey–Arens Theorem (see [29, Chap. IV, Sec. 3]) tells us that one can extend this strategy: the weak* topology is the weakest topology exhibiting this identification of E with its double dual E^{**}, but there are (possibly strictly stronger) topologies with the same property. So it is really only necessary to show that $\mathfrak{m}$ and $\mathfrak{q}$ are continuous w.r.t some topology η on E^* s.t. $(E^*, \eta)^* \simeq E$; in other words an admissible topology η. The strongest admissible topology is the Mackey topology w.r.t. the duality $\langle E^*, E\rangle$. And indeed, under the assumption of μ being Radon, $\mathfrak{m}$ and $\mathfrak{C}(g)$ are Mackey-continuous for every $g \in E^*$.

Theorem 3.9. *Let (E, τ) be a locally convex TVS and μ be a Gaussian measure on $\sigma(\tau)$ which is Radon. Then for any $g \in E^*$ the mappings $f \mapsto \mathfrak{m}(f)$ and $f \mapsto \big[\mathfrak{C}(g)\big](f) = \mathfrak{q}(g, f)$ are continuous w.r.t. the Mackey topology w.r.t. the duality $\langle E^*, E\rangle$ and thus representable as evaluation functionals.*

The assumption is satisfied when (E, τ) is separable Frechet and μ is defined on $\sigma(E; E^*)$.

Proof. [2, Thm. 3.2.1.] and its subsequent paragraphs show that for every $g \in E^*$ the operators $\mathfrak{m}$ and $\mathfrak{q}(g, -) = \mathfrak{C}(g)$ are continuous w.r.t. the Mackey topology on E^* w.r.t. the duality $\langle E^*, E\rangle$ and [29, Chap. IV, Sec. 3]) shows that this implies the result.

If (E, τ) is separable Frechet, then by Theorem 2.5 the weak and the Borel sigma-algebra coincide. Hence μ is Borel. Furthermore, since (E, τ) is Polish and μ is finite, μ is also Radon by [9, Chap. VIII, Thm. 1.16] and the assumption is satisfied. □

3.2.4 Regularity of the Covariance Operator

We have established that if E is a separable Banach space, then $\mathfrak{q}$ can be represented by a linear operator $\mathfrak{C}$ taking values in E. More can be said about its regularity.

Proposition 3.10. *Let $(E, \|\cdot\|_E)$ be a separable Banach space and μ a centred Gaussian measure on $\sigma(E; E^*)$. Then the covariance form $\mathfrak{q} : E^* \times E^* \to \mathbb{R}$ is bounded in the sense that there exists an $\alpha > 0$ s.t. for any $f, g \in E^*$*

$$|\mathfrak{q}(f,g)| \leq \alpha \|f\|_{E^*} \|g\|_{E^*}.$$

In particular, the covariance operator $\mathfrak{C}$ is compact with operator norm $\|\mathfrak{C}\| \leq \int_E \|x\|_E^2 \mathrm{d}\mu(x)$.

Proof. Let $f, g \in E^*$ be arbitrary. Then

$$\begin{aligned} |\mathfrak{q}(f,g)| = \left| \int_E f(x) g(x) \mathrm{d}\mu(x) \right| &\leq \int_E |f(x)||g(x)| \mathrm{d}\mu(x) \\ &\leq \int_E \|f\|_{E^*} \|g\|_{E^*} \|x\|_E^2 \mathrm{d}\mu(x) \\ &= \underbrace{\int_E \|x\|_E^2 \mathrm{d}\mu(x)}_{=:\alpha<\infty} \|f\|_{E^*} \|g\|_{E^*} \end{aligned} \tag{3.6}$$

where we used the triangle inequality for integrals and the standard bounds on the linear functional. To see the bound for $\|\mathfrak{C}\|$, note that

$$\|\mathfrak{C}(f)\|_E = \|\mathfrak{C}(f)\|_{E^{**}} = \sup_{\substack{g \in E^* \\ \|g\|_{E^*}=1}} |\mathfrak{q}(f,g)|$$

and apply the block of formulas above. Since E is separable Banach, the covariance operator $\mathfrak{C}$ takes values in $E \subseteq E^{**}$ by theorem 3.9. To see the compactness, let $(f_n)_{n\in\mathbb{N}}$ be a bounded sequence in E^*. Assume w.l.o.g. $\forall n \in \mathbb{N} : \|f_n\| \leq 1$. By the proof of Proposition 3.6 there exists a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ and an $f \in E^*$ s.t. $\|f_{n_k} - f\|_2 \to 0$. Thus

$$\begin{aligned} &\|\mathfrak{C}(f_{n_k}) - \mathfrak{C}(f)\|_E = \|\mathfrak{C}(f_{n_k} - f)\|_E = \|\mathfrak{C}(f_{n_k} - f)\|_{E^{**}} \\ &= \sup_{\substack{g \in E^* \\ \|g\|_{E^*}=1}} \left| \left[\mathfrak{C}(f_{n_k} - f)\right](g) \right| = \sup_{\substack{g \in E^* \\ \|g\|_{E^*}=1}} |\mathfrak{q}(f_{n_k} - f, g)| \\ &\leq \sup_{\substack{g \in E^* \\ \|g\|_{E^*}=1}} \|g\|_2 \|f_{n_k} - f\|_2 \leq \sup_{\substack{g \in E^* \\ \|g\|_{E^*}=1}} \|g\|_{E^*} \sqrt{\alpha} \|f_{n_k} - f\|_2 \leq \sqrt{\alpha} \|f_{n_k} - f\|_2 \to 0 \end{aligned}$$

where α comes from formula (3.6). Thus we have shown that the image of an arbitrary bounded sequence in E^* has a convergent subsequence in E, making $\mathfrak{C}$ compact. □

Proposition 3.11. *Let $(E, \langle\cdot,\cdot\rangle_E)$ be a separable Hilbert space and μ a centred Gaussian measure on $\sigma(E; E^*)$. Then $\mathfrak{C}$ is trace class with*

$$tr\ \mathfrak{C} = \int \|x\|_E^2 \ \mathrm{d}\mu(x)$$

Proof. Let $\{e_n\}_{n\in\mathbb{N}}$ be an orthonormal basis of E. Then

$$\int_E \|x\|_E^2 \,\mathrm{d}\mu(x) = \sum_{n=1}^{\infty} \int_E \langle x, e_n\rangle_E^2 \,\mathrm{d}\mu(x) = \sum_{n=1}^{\infty} \mathfrak{q}(e_n, e_n) = \sum_{n=1}^{\infty} \langle e_n, \mathfrak{C} e_n\rangle_E = \mathrm{tr}\ \mathfrak{C},$$

where in the first equality we used the Dominated Convergence Theorem with

$$\left|\sum_{n=1}^{N} \langle x, e_n\rangle_E^2\right| = \sum_{n=1}^{N} |\langle x, e_n\rangle|_E^2 \le \|x\|_E^2.$$

□

We may summarize the preceding two subsections by the following table

Measure & Space	Reg. of $\mathfrak{C}$	Representable	Bound	Reference
General on loc. co. TVS	weak* iff representable		-	Thm. 3.8
Radon on loc. co. TVS	Mackey	yes	-	Thm. 3.9
sep. Frechet	Mackey	yes	-	Thm. 3.9
sep. Banach	compact	yes	$\|\mathfrak{C}\| \le \|\big(\|\cdot\|_E\big)\|_2^2$	Prop. 3.10
sep. Hilbert	trace class	yes	$\mathrm{tr}\ \mathfrak{C} = \|\big(\|\cdot\|_E\big)\|_2^2$	Prop. 3.11

3.3 Cameron–Martin Space

In the rest of this thesis, unless explicitly stated otherwise, it is assumed that $\mathfrak{C}$ maps into $E \subseteq E^{**}$ instead of just $(E^*)'$. For the more general case see [2, Sec. 2.4.].

3.3.1 Generalities

Definition 3.12. Let (E,τ) be a locally convex TVS and μ a Gaussian measure on $\sigma(E;E^*)$ and denote the mapping $f \mapsto f - \mathfrak{m}(f) \in L^2(E,\mu)$ by $\mathtt{j}$.[7] Then the closure of the set

$$\mathtt{j}(E^*) = \big\{f - \mathfrak{m}(f) : f \in E^*\big\} \subseteq L^2(E,\mu) \tag{3.7}$$

with respect to the 2-norm on $L^2(E,\mu)$ is denoted by $\mathbf{K}(\mu)$. Formula (3.7) is well-defined by Proposition 3.6. Note that the values

[7] This coincides with the notation used in Proposition 3.6, since there the measure was assumed to be centred and thus $\mathfrak{m}(f) = 0$ for every $f \in E^*$.

$$
\begin{aligned}
\|\mathtt{j}(f)\|^2_{L^2(E,\mu)} &= \int_E \Big([\mathtt{j}(f)](x)\Big)^2 \mathrm{d}\mu(x) = \int_E \Big[f(x) - \mathfrak{m}(f)\Big]^2 \mathrm{d}\mu(x) \\
\|f\|^2_{\mathfrak{q}} &= \int_E \Big[f(x) - \mathfrak{m}(f)\Big]^2 \mathrm{d}\mu(x) \\
\mathrm{Var}\,\big[\mathtt{j}(f)\big] &= \int_E \Big[[\mathtt{j}(f)](x) - \underbrace{\mathbb{E}[\mathtt{j}(f)]}_{=\mathbb{E}(f)-\mathfrak{m}(f)=0}\Big]^2 \mathrm{d}\mu(x) = \int_E \Big[f(x) - \mathfrak{m}(f)\Big]^2 \mathrm{d}\mu(x)
\end{aligned}
$$

all coincide for every $f \in E^*$. Together with the inner product induced by $L^2(E,\mu)$ the space $\big(\mathbf{K}(\mu), \langle\cdot,\cdot\rangle_{L^2(E,\mu)}\big)$ becomes a Hilbert space, called the **reproducing kernel Hilbert space** of μ. The set of elements $x \in E$ s.t. $\mathrm{ev}_x : \big(\mathtt{j}(E^*), \langle\cdot,\cdot\rangle_{L^2(E,\mu)}\big) \to \mathbb{R}$ is a bounded linear functional is denoted $\mathbf{H}(\mu) \subseteq E$ and equipped with the operator norm, i.e. for any $x \in \mathbf{H}(\mu)$

$$
\|x\|_{\mathbf{H}(\mu)} := \sup_{\substack{f\in E^* \\ \mathtt{j}(f)\neq 0}} \left\{ \frac{\big|[\mathtt{j}(f)](x)\big|}{\sqrt{\langle \mathtt{j}(f), \mathtt{j}(f)\rangle_{L^2(E,\mu)}}} \right\} = \sup_{\substack{f\in E^* \\ \langle \mathtt{j}(f),\mathtt{j}(f)\rangle_{L^2(E,\mu)}=1}} \Big\{ \big|[\mathtt{j}(f)](x)\big| \Big\}.
$$

The space $\big(\mathbf{H}(\mu), \|\cdot\|_{\mathbf{H}(\mu)}\big)$ is called the **Cameron–Martin space** of μ in (E,τ). The Cameron–Martin space depends on E^* (and thus E and τ) and on $\langle\cdot,\cdot\rangle_{L^2(E,\mu)}$ (and thus μ).

Remark 3.13. Recall that elements in $L^2(E,\mu)$ are equivalence classes of functionals, not functionals per se. If μ is non-degenerate, then formula (3.7) really is an inclusion, i.e. the mapping $f \mapsto [f - \mathfrak{m}(f)]_{\mu-\text{a.s.}}$ is injective, where $[\ \cdot\ ]_{\mu-\text{a.s.}}$ denotes the equivalence class of functionals under the relation of μ-a.s. equivalence. The proof is essentially that of injectivity in Proposition 3.6.

The interesting behaviour (e.g. Proposition 3.15 and Proposition 3.21) that separates the infinite-dimensional setting from the finite-dimensional one is the fact that $\mathtt{j}(E^*)$ is infinite-dimensional. While this is certainly the case when $\dim E = \infty$ and μ is non-degenerate, it may also happen if $\dim E = \infty$ and μ is mildly degenerate. For example, if μ is the distribution of a Brownian bridge[8] on $C_0[0,1]$ tied down at 1, then the kernel of $\mathtt{j}$ is merely the linear span of $\{\mathrm{ev}_1\}$ and thus $\mathtt{j}(E^*)$ is infinite-dimensional. On the other end of the spectrum, if μ is very degenerate such as $\mu = \delta_0$ the Dirac measure at 0, then $\mathtt{j}(E^*) = \{0\}$ and the subsequent theory becomes trivial. We will make more remarks about degeneracy in appendix A.

Proposition 3.14. *Let (E,τ) be a locally convex TVS, μ a Gaussian measure on $\sigma(E;E^*)$, and $\Big(\mathbf{K}(\mu), \langle\cdot,\cdot\rangle_{L^2(E,\mu)}\Big)$ its reproducing kernel Hilbert space. Then*

(1) every functional $g \in \mathbf{K}(\mu)$ is centred Gaussian and has variance $\|g\|_2^2$,

(2) every functional $g \in \mathbf{K}(\mu)$ is μ-a.s. affine, and μ-a.s. linear if μ is centred,

[8] See subsection 4.4.

(3) if $f, g \in \mathbf{K}(\mu)$ are orthogonal w.r.t. the inner product on $L^2(E, \mu)$, then f and g are independent as random variables $(E, \mu) \to \mathbb{R}$.

Proof. (1) Let $g \in \mathbf{K}(\mu)$. Then there is a sequence $(f_n)_{n\in\mathbb{N}}$ in E^* s.t. $f_n - \mathfrak{m}(f_n) \to g$ in the L^2-norm and thus

$$\operatorname{Var}\left[\mathtt{j}(f_n)\right] = \left\|f_n - \mathfrak{m}(f_n)\right\|_{L^2}^2 \to \left\|g\right\|_{L^2}^2 = \operatorname{Var}(g).$$

Also, there is a subsequence $(f_{n_k} - \mathfrak{m}(f_{n_k}))_{k\in\mathbb{N}}$ s.t. $f_{n_k} - \mathfrak{m}(f_{n_k}) \to g$ μ-a.s. Thus, by virtue of the Dominated Convergence Theorem, we have for every $t \in \mathbb{R}$

$$\begin{aligned}\phi_g(t) = \widehat{\mu}(tg) &= \lim_{k\to\infty} \widehat{\mu}\Big[t\big(f_{n_k} - \mathfrak{m}(f_{n_k})\big)\Big] \\ &= \lim_{k\to\infty} \exp\left\{-\frac{1}{2}t^2 \operatorname{Var}\left[\mathtt{j}(f_{n_k})\right]\right\} = \exp\left\{-\frac{1}{2}t^2 \operatorname{Var}(g)\right\}.\end{aligned}$$

Hence g is Gaussian with distribution $\mathcal{N}(0, \operatorname{Var}(g))$.

(2) The existence of a μ-a.s. convergent subsequence in the proof above also shows that g is affine, and linear if μ is centred.

(3) Since $f, g \in \mathbf{K}(\mu)$ are centred, $\langle\cdot,\cdot\rangle_{L^2(E,\mu)}$ and thus

$$0 = \langle f, g\rangle_{L^2(E,\mu)} = \operatorname{Cov}(f, g).$$

Thus, f and g are uncorrelated and, since f and g are jointly Gaussian (see Remark 3.2), they are also independent. □

By definition, every element $g \in L^2(E, \mu)$ has finite variance equal to $\|g\|_2^2$. So what is remarkable about Proposition 3.14 is the first part, namely that all $g \in \mathbf{K}(\mu)$ are Gaussian. Note that not every element in $\mathbf{K}(\mu)$ is continuous as a functional $(E, \tau) \to \mathbb{R}$. The next theorem shows that passing from $\mathtt{j}(E^*)$ to its $\|\cdot\|_2$-closure is non-trivial.

Proposition 3.15. *[8, Cor. 4.17] Let $(E, \|\cdot\|_E)$ be a separable Banach space and μ a non-degenerate Gaussian measure. Then the normed space $(\jmath(E^*), \|\cdot\|_2)$ is not complete unless E is finite-dimensional.*

Proof. Since $\mathtt{j} : (E^*, \|\cdot\|_{E^*}) \to (E^*, \|\cdot\|_2)$ is bounded (by Proposition 3.6(2)), linear, and bijective, $\mathtt{j}$ is an isomorphism by the Open Mapping Theorem. But since $\mathtt{j}$ is also compact by Proposition 3.6, this implies that $B_1^{\|\cdot\|_2}(0)$ and thus $B_1^{\|\cdot\|_{E^*}}(0)$ is pre-compact, which implies that $(E^*, \|\cdot\|_{E^*})$, and thus E, is finite-dimensional. □

The covariance form $\mathfrak{q} : E^* \times E^* \to \mathbb{R}$ is L^2-continuous and the covariance operator $\mathfrak{C} : E^* \to E^{**}$ is L^2 - weak*-continuous[9]. Both are well-defined on equivalence classes and can thus be L^2-continuously extended to $\mathfrak{q} : \mathbf{K}(\mu) \times \mathbf{K}(\mu) \to \mathbb{R}$ and $\mathfrak{C} : \mathbf{K}(\mu) \to (E^*)'$. We will assume that the extension of $\mathfrak{C}$ also takes values in $E \subseteq E^{**}$, which is true when (E, τ) is weakly complete. This is the case, for example, when (E, τ) is a reflexive Banach space. The Cameron–Martin space can then be characterized as the image of this extended covariance operator as follows.

[9]Here we mean the weak* topology on E^{**}, which coincides with the weak topology on $E \subseteq E^{**}$.

Theorem 3.16. *Let (E, τ) be a locally convex TVS and μ a Gaussian measure on $\sigma(E; E^*)$. Then for any $h \in E$ we have*

$$h \in \mathbf{H}(\mu) \;\Leftrightarrow\; \exists! \overline{h} \in \mathbf{K}(\mu): \;\; ev_h = \mathfrak{C}(\overline{h}).$$

In this case $\|h\|_{\mathbf{H}(\mu)} = \left\|\overline{h}\right\|_{L^2(E,\mu)} = \left\|\overline{h}\right\|_{\mathfrak{q}}$.

Proof. "$\Rightarrow$": We show existence. The uniqueness then follows from the isometry. Let $h \in E$ with $\|h\|_{\mathbf{H}(\mu)} < \infty$. Then by definition ev_h defines a bounded linear functional on $\big(\mathtt{j}(E^*), \langle \cdot, \cdot \rangle_{L^2(E,\mu)}\big)$ and thus $\big(\mathbf{K}(\mu), \langle \cdot, \cdot \rangle_{L^2(E,\mu)}\big)$. Hence, by Riesz's Representation Theorem, there exists a unique $\overline{h} \in \mathbf{K}(\mu)$ s.t. for every $f \in E^*$

$$\mathrm{ev}_h\big(\mathtt{j}(f)\big) = \big[\mathtt{j}(f)\big](h) = \left\langle \overline{h}, \mathtt{j}(f) \right\rangle_{L^2(E,\mu)} = \int_E \overline{h}\mathtt{j}(f)\mathrm{d}\mu = \mathfrak{q}\left(\overline{h}, \mathtt{j}(f)\right) = \big[\mathfrak{C}(\overline{h})\big]\big(\mathtt{j}(f)\big).$$

Hence, by definition of $\mathfrak{C}$ and the density of $\mathtt{j}(E^*)$ in $\big(\mathbf{K}(\mu), \langle \cdot, \cdot \rangle_{L^2(E,\mu)}\big)$, $h = \mathfrak{C}(\overline{h})$.

"$\Leftarrow$": Let $h = \mathfrak{C}(\overline{h})$ for some $\overline{h} \in \mathbf{K}(\mu)$. Then

$$\begin{aligned} \|h\|_{\mathbf{H}(\mu)} &= \sup\left\{ \left|\mathrm{ev}_h\big(\mathtt{j}(f)\big)\right| : f \in E^*, \langle \mathtt{j}(f), \mathtt{j}(f) \rangle_{L^2(E,\mu)} = 1 \right\} \\ &= \sup\left\{ \left| \left\langle \overline{h}, \mathtt{j}(f) \right\rangle_{L^2(E,\mu)} \right| : f \in E^*, \langle \mathtt{j}(f), \mathtt{j}(f) \rangle_{L^2(E,\mu)} = 1 \right\} \\ &\geq \left\langle \overline{h}, \frac{\overline{h}}{\left\|\overline{h}\right\|_{L^2(E,\mu)}} \right\rangle_{L^2(E,\mu)} \\ &= \left\|\overline{h}\right\|_{L^2(E,\mu)}. \end{aligned} \tag{3.8}$$

This gives $\|h\|_{\mathbf{H}(\mu)} \geq \left\|\overline{h}\right\|_{L^2(E,\mu)}$. The Cauchy–Schwarz inequality applied to the inner product in formula (3.8) also provides $\left\|\overline{h}\right\|_{L^2(E,\mu)}$ as an upper bound, which shows that $\|h\|_{\mathbf{H}(\mu)}$ is finite and also shows the isometry (which proves the uniqueness of $\overline{h}$ for a given h). □

Thus $\mathfrak{C}$ defines a linear, isometric surjection, i.e. an isometric isomorphism $\mathfrak{C}: \big(\mathbf{K}(\mu), \langle \cdot, \cdot \rangle_{L^2(E,\mu)}\big) \to \big(\mathbf{H}(\mu), \|\cdot\|_{\mathbf{H}(\mu)}\big)$, which turns $\big(\mathbf{H}(\mu), \|\cdot\|_{\mathbf{H}(\mu)}\big)$ into a Hilbert space via

$$\langle h, k \rangle_{\mathbf{H}(\mu)} = \left\langle \overline{h}, \overline{k} \right\rangle_{L^2(E,\mu)} = \mathfrak{q}\left(\overline{h}, \overline{k}\right). \tag{3.9}$$

The above can be summarized by the following diagram.

$$\begin{tikzcd} (E,\tau)^* \arrow[r, hook, "\mathtt{j}"] & \big(\mathbf{K}(\mu), \langle \cdot, \cdot \rangle_{L^2(E,\mu)}\big) \subseteq L^2(E,\mu) \arrow[dl, bend right, "\mathfrak{C}"'] \\ \big(\mathbf{H}(\mu), \langle \cdot, \cdot \rangle_{\mathbf{H}(\mu)}\big) \arrow[ur, bend right, "\overline{(\cdot)}"'] \arrow[r, hook, "\mathtt{i}"] & (E,\tau) \end{tikzcd} \qquad \begin{tikzcd} f \arrow[r, mapsto, "\mathtt{j}"] & \big(\mathtt{j}(f)\big) = \overline{\mathfrak{C}\big(\mathtt{j}(f)\big)} \arrow[dl, mapsto, bend right, "\mathfrak{C}"'] \\ \left\langle \cdot, \mathfrak{C}\big(\mathtt{j}(f)\big) \right\rangle_{\mathbf{H}(\mu)} = \mathfrak{C}\big(\mathtt{j}(f)\big) \arrow[ur, mapsto, bend right, "\overline{(\cdot)}"'] \arrow[r, mapsto, "\mathtt{i}"] & \mathfrak{C}\big(\mathtt{j}(f)\big) \end{tikzcd}$$

3.3.2 Special Case: Finite-Dimensional Space

Consider what this means for for the finite-dimensional case. Let $\mu = \mathcal{N}(0, \Sigma)$ be a centred, non-degenerate Gaussian measure on $\mathbb{R}^n$. Then the assumption at the beginning of the chapter is satisfied, and since Σ is assumed to be non-degenerate, the diagram takes the form

$$\begin{tikzcd}
(\mathbb{R}^n)^* \arrow[r, hook, "\mathtt{j}"] & \mathtt{j}\left[(\mathbb{R}^n)^*\right] \subseteq L^2(\mathbb{R}^n, \mu) \arrow[dl, "\mathfrak{C}"'] \\
\mathbb{R}^n \arrow[r, hook, "\mathtt{i}"] & \mathbb{R}^n
\end{tikzcd}
\qquad
\begin{tikzcd}
\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \cdot \right\rangle_{\mathbb{R}^n} \arrow[r, mapsto, "\mathtt{j}"] & \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \cdot \right\rangle_{\mathbb{R}^n} \arrow[dl, mapsto, "\mathfrak{C}"'] \\
\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \arrow[r, mapsto, "\mathtt{i}"] & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}
\end{tikzcd}$$

The bilinear forms $\mathfrak{q}$ and $\langle \cdot, \cdot \rangle_{\mathbf{H}(\mu)}$ can be identified as follows. As shown in Remark 2.9 (or as can be seen by comparing the finite-dimensional formula of the characteristic function of $\mathcal{N}(0, \Sigma)$ with Theorem 3.4), the covariance form is representable by Σ in the sense that

$$\mathfrak{q}\big(\langle x, \cdot \rangle_{\mathbb{R}^n}, \langle y, \cdot \rangle_{\mathbb{R}^n}\big) = \langle x, \Sigma y \rangle_{\mathbb{R}^n}, \quad \forall x, y \in \mathbb{R}^n.$$

Let $\langle x, \cdot \rangle_{\mathbb{R}^n}$ be an element of $\mathtt{j}\left[(\mathbb{R}^n)^*\right]$. Then applying $\mathfrak{C}$ needs to yield an element of $\mathbb{R}^n$ s.t. for any functional $\langle y, \cdot \rangle_{\mathbb{R}^n}$

$$\Big\langle y, \mathfrak{C}\langle x, \cdot \rangle_{\mathbb{R}^n} \Big\rangle_{\mathbb{R}^n} = \mathfrak{q}\big(\langle y, \cdot \rangle_{\mathbb{R}^n}, \langle x, \cdot \rangle_{\mathbb{R}^n}\big) = \langle y, \Sigma x \rangle_{\mathbb{R}^n}.$$

Thus $\mathfrak{C}\langle x, \cdot \rangle_{\mathbb{R}^n} = \Sigma x$ and $(\bar{\cdot})$ is given by $x \mapsto \langle \Sigma^{-1} x, \cdot \rangle_{\mathbb{R}^n}$. Hence for any $x, y \in \mathbb{R}^n$

$$\langle x, y \rangle_{\mathbf{H}(\mu)} = \mathfrak{q}\Big(\left\langle \Sigma^{-1} x, \cdot \right\rangle_{\mathbb{R}^n}, \left\langle \Sigma^{-1} y, \cdot \right\rangle_{\mathbb{R}^n} \Big) = \left\langle \Sigma^{-1} x, \Sigma \Sigma^{-1} y \right\rangle_{\mathbb{R}^n} = \left\langle x, \Sigma^{-1} y \right\rangle_{\mathbb{R}^n}.$$

Note that $\langle \cdot, \cdot \rangle_{\mathbf{H}(\mu)}$ is the bilinear form in the Gaussian density, i.e. we may write $\mathrm{d}\mu$ as

$$\frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left\{ -\frac{1}{2} \|x\|^2_{\mathbf{H}(\mu)} \right\} \, \mathrm{d}\lambda^n(x), \quad x \in \mathbb{R}^n, \tag{3.10}$$

where λ^n is the n-dimensional Lebesgue measure on $\mathbb{R}^n$. On $\mathbb{R}^n$ this is completely rigorous, and, despite the fact that formula (3.10) does not make any sense on an infinite-dimensional TVS, it provides a good heuristic. Figure 4 is a plot of the unit circle of the $\mathbf{H}(\mu)$-norm for a centred Gaussian measure on $\mathbb{R}^2$ with covariance matrix $\begin{pmatrix} 3 & 1 \\ 1 & 0.5 \end{pmatrix}$. It may provide some intuition for how the Cameron–Martin space looks like in general.

3.3.3 Special Case: Classical Wiener Space

In the case of the classical Wiener space (see chapter 4.1) the necessary assumptions are also satisfied and the diagram takes the form

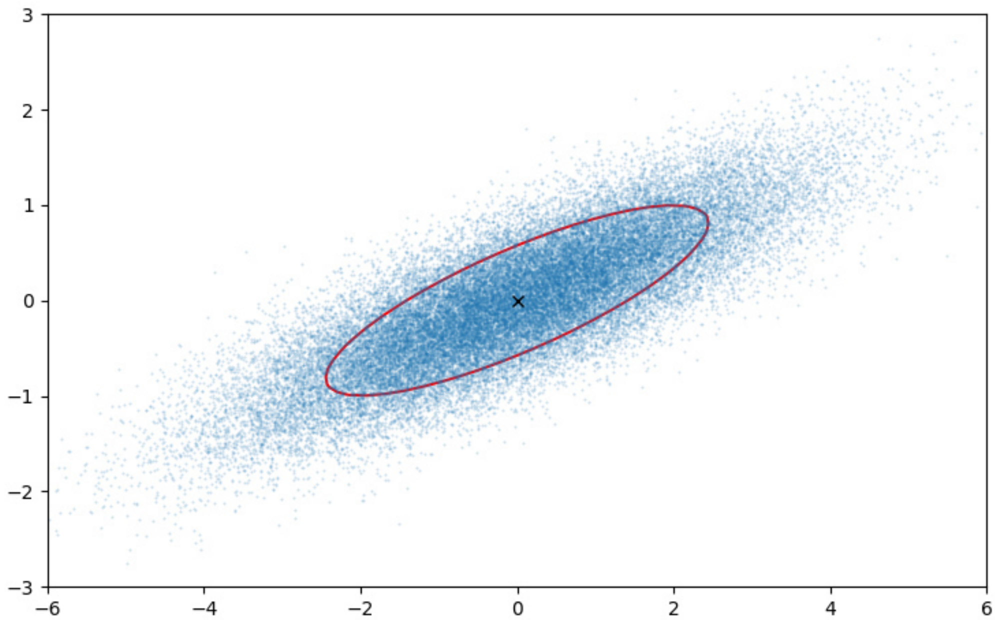

Figure 4: Sample points and the set $\{x \in \mathbb{R}^2 : \|x\|_{\mathbf{H}(\mu)} = 1\}$ for a centred Gaussian measure on $\mathbb{R}^2$ with covariance matrix $\left(\begin{smallmatrix} 3 & 1 \\ 1 & 0.5 \end{smallmatrix}\right)$

$$\begin{tikzcd}
\mathbf{M}_0([0,1]) \arrow[r, hook, "\mathtt{j}"] & (\mathbf{K}(\mu), \mathfrak{q}) \subseteq L^2(C_0[0,1], \mu) \arrow[dl, bend right, "\mathfrak{C}"'] \\
(H_0^1, \langle \cdot, \cdot \rangle_H) \arrow[ur, bend right, "\overline{(\cdot)}"'] \arrow[r, hook, "\mathtt{i}"] & C_0[0,1]
\end{tikzcd}
\qquad
\begin{tikzcd}
\int_0^1 \cdot \, \mathrm{d}\delta_s \arrow[r, mapsto, "\mathtt{j}"] & \int_0^1 \cdot \, \mathrm{d}\delta_s = \int_0^1 \Psi_s'(t) \, \mathrm{d} \cdot (t) = \overline{\Psi_s} \arrow[dl, mapsto, bend right, "\mathfrak{C}"'] \\
\Psi_s(t) = s \wedge t \arrow[ur, mapsto, bend right, "\overline{(\cdot)}"'] \arrow[r, mapsto, "\mathtt{i}"] & \Psi_s(t) = s \wedge t
\end{tikzcd}$$

By the Riesz–Markov–Kakutani Representation Theorem, the continuous dual space of $C[0,1]$ can be identified with the space $\mathbf{M}([0,1])$, consisting of the finite signed Borel measures on $[0,1]$. Consequently, the continuous dual space of the space of continuous functions on $[0,1]$ s.t. $x(0) = 0$ is $\mathbf{M}_0([0,1])$, defined as the quotient of $\mathbf{M}([0,1])$ by the span of $\{\delta_0\}$, with the action of $\nu \in \mathbf{M}_0([0,1])$ on an $x \in C_0[0,1]$ given by $x \mapsto \int_0^1 x(t) \mathrm{d}\nu(t)$. We will consider the action of $\mathfrak{C}$ on $D := \{\delta_t : t \in [0,1]\}$, as this uniquely determines $\mathfrak{C}$ on $(\mathbf{K}(\mu), \mathfrak{q})$ since the linear span of D lies dense in $(\mathbf{K}(\mu), \mathfrak{q})$.

By virtue of Fubini's Theorem, the covariance form of a continuous and centred Gaussian process $(X_t)_{t \in [0,1]}$, defined on some probability space $(\Omega, \mathbb{P})$, can be written as

$$
\begin{aligned}
\mathfrak{q}(\nu_1,\nu_2) &= \int_{C[0,1]} \left[\int_0^1 x(t)\mathrm{d}\nu_1(t)\right]\left[\int_0^1 x(s)\mathrm{d}\nu_2(s)\right]\mathrm{d}\mu(x) \\
&= \int_{C[0,1]}\int_0^1\int_0^1 \mathrm{ev}_s(x)\,\mathrm{ev}_t(x)\mathrm{d}\nu_2(s)\mathrm{d}\nu_1(t)\mathrm{d}\mu(x) \\
&= \int_0^1\int_0^1\int_{C[0,1]} \mathrm{ev}_s(x)\,\mathrm{ev}_t(x)\mathrm{d}\mu(x)\mathrm{d}\nu_2(s)\mathrm{d}\nu_1(t) \\
&= \int_0^1\int_0^1\int_{C[0,1]} X_sX_t\mathrm{d}\mathbb{P}\mathrm{d}\nu_2(s)\mathrm{d}\nu_1(t) \\
&= \int_0^1\int_0^1 \mathrm{Cov}(X_s,X_t)\mathrm{d}\nu_2(s)\mathrm{d}\nu_1(t),
\end{aligned}
$$

where $\mathrm{Cov}(X_s,X_t)$ denotes the **covariance structure** of $(X_t)_{t\in[0,1]}$. In the case of BM this is $\mathrm{Cov}(X_s,X_t)=s\wedge t$ for every $s,t\in[0,1]$.

Let $\nu\in\mathbf{M}_0([0,1])$ be arbitrary. Since D lies weak*-dense in $\mathbf{M}_0([0,1])$ and $\mathbf{M}_0([0,1])$ separates points, we may identify $\mathfrak{C}(\nu)$ by testing it against Dirac measures. Thus by the above derivation, $\mathfrak{C}(\nu)$ needs to be an element of $C_0[0,1]$ s.t. for every $t\in[0,1]$

$$
\begin{aligned}
[\mathfrak{C}(\nu)](t) &= \int_0^1 [\mathfrak{C}(\nu)](s)\ \mathrm{d}\delta_t(s) = \mathrm{ev}_{\mathfrak{C}(\nu)}(\delta_t) = [\mathfrak{C}(\nu)](\delta_t) = \mathfrak{q}(\nu,\delta_t) \\
&= \int_0^1\int_0^1 \mathrm{Cov}(X_s,X_u)\ \mathrm{d}\delta_t(s)\ \mathrm{d}\nu(u) = \int_0^1 \mathrm{Cov}(X_t,X_u)\ \mathrm{d}\nu(u).
\end{aligned}
$$

In particular, the images of the Dirac measures are

$$[\mathfrak{C}(\delta_s)](t) = \mathrm{Cov}(X_t,X_s) = t\wedge s =: \Psi_s(t).$$

See also the diagram for the classical Wiener space. According to Theorem 3.16 the $\mathbf{H}(\mu)$-norm for images of D is given by

$$\langle \Psi_s,\Psi_t\rangle_{\mathbf{H}(\mu)} = \mathfrak{q}(\delta_s,\delta_t) = s\wedge t,$$

which coincides with the Dirichlet form

$$\int_0^1 \Psi_s'(u)\Psi_t'(u)\mathrm{d}\lambda(u) = \langle\Psi_s,\Psi_t\rangle_{H_0^1[0,1]}.$$

Since $\mathfrak{C}$ is a linear isomorphism, this means that $\mathbf{H}(\mu)$ is the closure of $\{\Psi_t : t\in[0,1]\}$ under the above norm, which is the first Hilbert–Sobolev space $H_0^1[0,1]$ of functions with $x(0)=0$. That is

$$H_0^1[0,1] = \left\{h\in C_0[0,1] : \exists h'\in L^2[0,1] \text{ s.t. } \forall t\in[0,1] : h(t) = \int_0^t h'(s)\mathrm{d}s\right\}.$$

Following the heuristic introduced in the previous section we may formally (!) see the classical Wiener measure as

$$\frac{1}{\beta}\exp\left\{-\frac{1}{2}\|x\|^2_{H_0^1[0,1]}\right\}\mathcal{D}x, \quad x \in C_0[0,1],$$

where $\mathcal{D}$ is the "infinite-dimensional Lebesgue measure on $C_0[0,1]$" and β is a normalizing constant. Again, $\|x\|^2_{H_0^1[0,1]}$ does not make any sense for an element $x \in C_0[0,1] \setminus H_0^1[0,1]$. It is also worthwhile to study the inverse map to $\mathfrak{C}$. As before, we only need to consider its action on

$$\mathfrak{C}\big\{\delta_t : t \in [0,1]\big\} = \big\{\Psi_t : t \in [0,1]\big\} \subseteq C_0[0,1].$$

For every $t \in [0,1]$ the functional $\overline{\Psi_t}$ is defined for μ-a.e. $x \in C_0[0,1]$ and it coincides with $\mathrm{ev}_t = \int_0^1 \cdot \, \mathrm{d}\delta_t$. For a general $h \in H_0^1[0,1]$ there exists an approximation of the form

$$\sum_{n=1}^{N} \alpha_n \Big(\Psi_{t_n} - \Psi_{t_{n-1}}\Big) \to h$$

in the $H_0^1[0,1]$-norm with $\alpha_n \in \mathbb{R}$ and hence

$$\sum_{n=1}^{N} \alpha_n \Big(\overline{\Psi}_{t_n} - \overline{\Psi}_{t_{n-1}}\Big) \to \overline{h}$$

in the L^2-norm. Thus there exists a subsequence s.t.

$$\sum_{k=1}^{N_k} \alpha_{n_k} \Big(\overline{\Psi}_{t_{n_k}} - \overline{\Psi}_{t_{n_k-1}}\Big) \to \overline{h} \quad \mu-\text{a.s.},$$

and thus for μ-a.e. $x \in C_0[0,1]$

$$\overline{h}(x) = \lim_{k\to\infty} \sum_{k=1}^{N_k} \alpha_{n_k} \Big(\overline{\Psi}_{t_n}(x) - \overline{\Psi}_{t_{n-1}}(x)\Big) = \lim_{k\to\infty} \sum_{k=1}^{N_k} \alpha_{n_k} \big(x(t_n) - x(t_{n-1})\big).$$

In other words, $\mathbb{P}$-a.s.

$$\overline{h}\big((B_t)_{t\in[0,1]}\big) = \lim_{k\to\infty} \sum_{k=1}^{N_k} \alpha_{n_k} \big(B_{t_n} - B_{t_{n-1}}\big),$$

which is nothing but the Paley–Wiener stochastic integral associated to $h \in H_0^1[0,1]$, where $(B_t)_{t\in[0,1]}$ is a standard Brownian motion defined on some probability space $(\Omega, \mathbb{P})$. It is usually denoted

$$W(h) := \int_0^1 h'(t)\mathrm{d}B_t.$$

The preceding also entails the Itô-isometry in the form of equation (3.9) as

$$\|W(h)\|_{L^2(\Omega,\mathbb{P})} = \left\|\int_0^1 \dot{h}(t)\mathrm{d}B_t\right\|_{L^2(\Omega,\mathbb{P})} = \|h\|_{H_0^1[0,1]} = \|h'\|_{L^2[0,1]}.$$

3.4 Cameron–Martin Theorems

As alluded to in section 1.3, a Gaussian measure on $\mathbb{R}^n$ is quasi-invariant w.r.t. translation in any direction. That is, for $\mu = \mathcal{N}(0, \Sigma)$ and any $y \in \mathbb{R}^n$ the measures μ and μ_y are equivalent, meaning that both Radon–Nikodým derivatives $\frac{\mathrm{d}\mu_y}{\mathrm{d}\mu}$ and $\frac{\mathrm{d}\mu}{\mathrm{d}\mu_y}$ exist. Moreover, they can be computed explicitly: For any $A \in \mathcal{B}(\mathbb{R}^n)$

$$\begin{aligned}
\mu_y(A) &= \mu(A-y) \\
&= \int_{A-y} \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left\{ -\frac{1}{2} \left\langle x, \Sigma^{-1} x \right\rangle_{\mathbb{R}^n} \right\} \mathrm{d}\lambda^n(x) \\
&= \int_{A} \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left\{ -\frac{1}{2} \left\langle x-y, \Sigma^{-1} x-y \right\rangle_{\mathbb{R}^n} \right\} \mathrm{d}\lambda^n(x) \\
&= \int_{A} \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left\{ -\frac{1}{2} \left\langle x, \Sigma^{-1} x \right\rangle_{\mathbb{R}^n} + \left\langle x, \Sigma^{-1} y \right\rangle_{\mathbb{R}^n} - \frac{1}{2} \left\langle y, \Sigma^{-1} y \right\rangle_{\mathbb{R}^n} \right\} \mathrm{d}\lambda^n(x) \\
&= \int_{A} \exp\left\{ \left\langle x, \Sigma^{-1} y \right\rangle_{\mathbb{R}^n} - \frac{1}{2} \left\langle y, \Sigma^{-1} y \right\rangle_{\mathbb{R}^n} \right\} \mathrm{d}\mu(x)
\end{aligned}$$

and thus

$$\frac{\mathrm{d}\mu_y}{\mathrm{d}\mu}(x) = \exp\left\{ \left\langle x, \Sigma^{-1} y \right\rangle_{\mathbb{R}^n} - \frac{1}{2} \left\langle y, \Sigma^{-1} y \right\rangle_{\mathbb{R}^n} \right\}, \quad x \in \mathbb{R}^n. \tag{3.11}$$

As opposed to translation-invariance, this formula can be generalized to general locally convex TVS, albeit in the infinite-dimensional case, not every $y \in E$ is admissible, but only those y which lie the Cameron–Martin space. The following theorem will thus (partially) resolve the problem encountered in section 1.1.

Theorem 3.17. *(Cameron–Martin I, [2, Prop. 2.4.2.]) Let (E, τ) be a locally convex TVS, μ a Gaussian measure on $\sigma(E; E^*)$, and $h \in \mathbf{H}(\mu)$. Then $\mu_h := \mu(\cdot - h)$ is Gaussian and equivalent to μ, and the Radon–Nikodým derivative is given by*

$$\frac{\mathrm{d}\mu_h}{\mathrm{d}\mu}(x) = \exp\left\{ \overline{h}(x) - \frac{1}{2} \|h\|^2_{\mathbf{H}(\mu)} \right\}, \quad x \in E, \tag{3.12}$$

where $\overline{(\cdot)}$ is the inverse mapping of the covariance operator $\mathfrak{C}$ (see Theorem 3.16).

Proof. By Fernique's Theorem $\exp\left(|\overline{h}|\right)$ is integrable and hence (3.12) gives a finite and non-negative measure. We want to show that for any $f \in E^*$

$$\widehat{\mu_h}(f) = \int_E \exp\left\{ i f(x) \right\} \exp\left\{ \overline{h}(x) - \frac{1}{2} \|h\|^2_{\mathbf{H}(\mu)} \right\} \mathrm{d}\mu(x), \tag{3.13}$$

where the latter is the Fourier transform of a measure given by the density (3.12). Then the Fourier Uniqueness Theorem 2.15 gives the result. Firstly, for any $f \in E^*$

$$\begin{aligned}
\widehat{\mu_h}(f) &= \int_E \exp\Big\{if(x)\Big\}\,\mathrm{d}\mu_h(x)\\
&= \int_E \exp\Big\{if(x+h)\Big\}\,\mathrm{d}\mu(x)\\
&= \int_E \exp\Big\{i\big[f(x)+f(h)\big]\Big\}\,\mathrm{d}\mu(x)\\
&= \exp\Big\{i\mathfrak{m}\Big[f+f(h)\Big]-\frac{1}{2}\mathfrak{q}\Big[f(x)+f(h),f(x)+f(h)\Big]\Big\}\\
&= \exp\Big\{i\mathfrak{m}(f)+if(h)-\frac{1}{2}\mathfrak{q}\Big[f+f(h),f+f(h)\Big]\Big\}\\
&= \exp\Big\{i\mathfrak{m}(f)+if(h)-\frac{1}{2}\Big[\mathfrak{q}(f,f)+0+0\Big]\Big\}\\
&= \exp\Big\{i\mathfrak{m}(f)+i\langle\overline{h},f\rangle_{L^2(E,\mu)}-\frac{1}{2}\mathfrak{q}(f,f)\Big\}. \qquad (3.14)
\end{aligned}$$

Now define the function

$$\rho_f(z)=\exp\Big\{i\mathfrak{m}(f)-\frac{1}{2}\|h\|^2_{\mathbf{H}(\mu)}\Big\}\int_E \exp\Big\{i\Big[f(x)-\mathfrak{m}(f)-z\overline{h}(x)\Big]\Big\}\,\mathrm{d}\mu(x) \qquad (3.15)$$

of a real variable z. Then, since $f-\mathfrak{m}(f)-z\overline{h}$ lies in $\mathbf{K}(\mu)$, and is thus centred Gaussian by Proposition 3.14, the integral of (3.15) becomes

$$\begin{aligned}
&\exp\Big\{-\frac{1}{2}\mathfrak{q}\Big[f-\mathfrak{m}(f)-z\overline{h},f-\mathfrak{m}(f)-z\overline{h}\Big]\Big\}\\
&=\exp\Big\{-\frac{1}{2}\mathfrak{q}\Big[f-\mathfrak{m}(f),f-\mathfrak{m}(f)\Big]-\frac{1}{2}z^2\mathfrak{q}\left(\overline{h},\overline{h}\right)+z\mathfrak{q}\left(\overline{h},f-\mathfrak{m}(f)\right)\Big\}\\
&=\exp\Big\{-\frac{1}{2}\mathfrak{q}(f,f)-\frac{1}{2}z^2\|\overline{h}\|^2_{L^2(E,\mu)}+z\langle\overline{h},f\rangle_{L^2(E,\mu)}\Big\}\\
&=\exp\Big\{-\frac{1}{2}\mathfrak{q}(f,f)-\frac{1}{2}z^2\|h\|^2_{\mathbf{H}(\mu)}+z\langle\overline{h},f\rangle_{L^2(E,\mu)}\Big\},
\end{aligned}$$

which gives

$$\rho_f(z)=\exp\Big\{i\mathfrak{m}(f)-\frac{1}{2}\|h\|^2_{\mathbf{H}(\mu)}\Big\}\exp\Big\{-\frac{1}{2}\mathfrak{q}(f,f)-\frac{1}{2}z^2\|h\|^2_{\mathbf{H}(\mu)}+z\langle\overline{h},f\rangle_{L^2(E,\mu)}\Big\}. \qquad (3.16)$$

Since ρ_f is analytic with radius of convergence $r=\infty$, there exists a holomorphic extension to all of $\mathbb{C}$. Now, on the one hand, for arbitrary $z_n\to i$, using the Bounded Convergence Theorem with

$$\left|\exp\Big\{i\Big[f(x)-\mathfrak{m}(f)-z_n\overline{h}(x)\Big]\Big\}\right|\le\exp\Big\{\max_{n\in\mathbb{N}}|iz_n|\,\left|\overline{h}(x)\right|\Big\}\quad,\forall x\in E, \qquad (3.17)$$

the expression in (3.15) converges to the right hand side of equation (3.13). On the other hand, the right hand side of equation (3.16) converges to formula (3.14), which is the left hand side of equation (3.13). Thus for any $f \in E^*$ the sequence $\rho_f(z_n)$ converges to both the left and right hand side of equation (3.13), which means that they are equal. This is what was to be shown.

To see not only absolute continuity of μ_h w.r.t. μ, but also equivalence, note that $\mu = (\mu_h)_{-h}$ and apply the same argument as above. □

The formula for the Radon–Nikodým density (3.12) is called the (abstract) **Cameron–Martin formula**.

Lemma 3.18. *Let μ be a Gaussian measure on $\mathbb{R}^n$ and $x \in \mathbb{R}^n$ arbitrary. Then*

$$2 - 2\exp\left\{ -\frac{1}{8}\left\langle x, \Sigma^{-1}x\right\rangle_{\mathbb{R}^n} \right\} \leq \|\mu - \mu_h\|_{TV}$$

where $\|\cdot\|_{TV}$ denotes the total variation norm.

Proof. See [2, Lem. 2.4.4.]. □

Theorem 3.19. *(Cameron–Martin II, [2, Thm. 2.4.5.(i)]) Let (E, τ) be a locally convex TVS, μ a Gaussian measure on $\sigma(E; E^*)$, and $h \in E$ s.t. $\|h\|_{\mathbf{H}(\mu)} = \infty$. Then μ_h and μ are mutually singular.*

Proof. By assumption of $\|h\|_{\mathbf{H}(\mu)} = \infty$, for every $n \in \mathbb{N}$ there exists an $f_n \in E^*$ s.t. $\mathfrak{q}(f_n, f_n) = 1$ and $f_n(h) \geq n$. Then

$$\begin{aligned}
\|\mu_h - \mu\|_{\mathrm{TV}} &= \sup\Big\{ |\mu_h(A) - \mu(A)| : A \in \sigma(E; E^*) \Big\} \\
&\geq \sup\left\{ \left| \left[\mu \circ f_n^{-1}\right]_{f_n(h)}(B) - \left[\mu \circ f_n^{-1}\right](B) \right| : B \in \mathcal{B}(\mathbb{R}) \right\} \\
&= \left\| \left[\mu \circ f_n^{-1}\right]_{f_n(h)} - \left[\mu \circ f_n^{-1}\right] \right\|_{\mathrm{TV}}
\end{aligned}$$

By Lemma 3.18, for every $n \in \mathbb{N}$

$$\left\| \left[\mu \circ f_n^{-1}\right]_{f_n(h)} - \left[\mu \circ f_n^{-1}\right] \right\|_{\mathrm{TV}} \geq 2 - 2\exp\left\{ -\frac{1}{8}|f(h)|^2 \right\} \geq 2 - 2\exp\left\{ -\frac{1}{8}n^2 \right\},$$

so $\|\mu_h - \mu\|_{\mathrm{TV}} = 2$ and thus μ_h and μ are mutually singular. □

The upshot of Theorem 3.17 and Theorem 3.19 is that shifting a Gaussian measure μ along an element $x \in E$ yields an equivalent measure if and only if $x \in \mathbf{H}(\mu)$. In that case, the Radon–Nikodým derivative can be computed explicitly. We conclude this chapter with two interesting properties of $\big(\mathbf{H}(\mu), \|\cdot\|_{\mathbf{H}(\mu)}\big)$. The first one is that if E is a separable Banach space, then the topology induced by the Cameron–Martin norm is "a lot" stronger than the norm-topology. This manifests itself in the following proposition.

Proposition 3.20. *Let $(E, \|\cdot\|_E)$ be a separable Banach space, μ a Gaussian measure on $\sigma(E; E^*)$, and $\mathbf{H}(\mu)$ its Cameron–Martin space. Then*

$$B_1^{\mathbf{H}(\mu)}(0) := \{h \in E : \|h\|_{\mathbf{H}(\mu)} \leq 1\} \subseteq E$$

is compact in the norm-topology of E. A similar result is generally true when E is assumed to be only locally convex and the norm-topology is replaced by the weak topology.

Proof. See [8, p. 16, bottom] for the claim for separable Banach spaces and [2, Prop. 2.4.6.] for the claim regarding general locally convex TVSs. □

The second one is that, despite its importance in the theory, the Cameron–Martin space is quite small as far as the measure itself is concerned.

Proposition 3.21. *([2, Prop. 2.4.7.]) Let (E, τ) be a locally convex TVS, μ a centred Gaussian measure on $\sigma(E; E^*)$, and $\mathbf{H}(\mu)$ its Cameron–Martin space. Then*

$$\mathbf{H}(\mu) = \bigcap_{\substack{L \text{ linear subspace} \\ L \in \sigma(E; E^*) \\ \mu(L) = 1}} L,$$

and, unless $\mathbf{K}(\mu)$ has finite dimension,

$$\mu\big[\mathbf{H}(\mu)\big] = 0.$$

Proof. $\subseteq$: Let $h \in \mathbf{H}(\mu)$ be arbitrary and let L be a measurable linear subspace with $\mu(L) = 1$. Then $\mu(L + h) = \mu_h(L) = 1$ by Theorem 3.17. Thus $h \in L$, as otherwise

$$\mu(E) \geq \mu(L) + \mu(L + h) \geq 2.$$

$\supseteq$: Let $h \notin \mathbf{H}(\mu)$ be arbitrary. Then by definition there exist a sequence $(f_n)_{n\in\mathbb{N}}$ in E^* s.t. $\langle \mathtt{j}(f_n), \mathtt{j}(f_n)\rangle_{L^2(E,\mu)} = 1$, but $\big[\mathtt{j}(f_n)\big](h) \geq n$. Now notice that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \int_E \big|\big[\mathtt{j}(f_n)\big](x)\big|^2 \, \mathrm{d}\mu(x) = \sum_{n=1}^{\infty} \frac{1}{n^2} \langle \mathtt{j}(f_n), \mathtt{j}(f_n)\rangle_{L^2(E,\mu)} = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Thus via the Monotone Convergence Theorem the first term equals

$$\int_E \sum_{n=1}^{\infty} \frac{1}{n^2} \big|\big[\mathtt{j}(f_n)\big](x)\big|^2 \, \mathrm{d}\mu < \infty,$$

and we conclude that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \big|\big[\mathtt{j}(f_n)\big](x)\big|^2 < \infty \quad \text{for } \mu-\text{a.e. } x \in E.$$

Thus the subspace

$$L := \Big\{x \in E : \sum_{n=1}^{\infty} \frac{1}{n^2} \big|\big[\mathtt{j}(f_n)\big](x)\big|^2 < \infty\Big\} = \bigcup_{M\in\mathbb{N}} \bigcap_{N\in\mathbb{N}} \Big\{x \in E : \sum_{n=1}^{N} \frac{1}{n^2} \big|\big[\mathtt{j}(f_n)\big](x)\big|^2 < M\Big\}$$

is measurable and has full measure. However, by construction, $h \notin L$. To see the second claim, let $\mathbf{K}(\mu)$ have infinite dimension. Then there exists an infinite orthonormal set $(\mathtt{j}(f_n))_{n\in\mathbb{N}}$. By Proposition 3.14, $(\mathtt{j}(f_n))_{n\in\mathbb{N}}$ is a sequence of iid random variables $(E,\mu) \to \mathbb{R}$ s.t. $f_1 \sim \mathcal{N}(0,1)$. Thus for μ-a.e. $x \in E$ the sequence $\big([\mathtt{j}(f_n)](x)\big)_{n\in\mathbb{N}}$ is unbounded and in particular

$$\Big\{x \in E : \sum_{n=1}^{\infty} \big|[\mathtt{j}(f_n)](x)\big|^2 < \infty\Big\}$$

has measure 0, but contains $\mathbf{H}(\mu)$ since for every $h \in \mathbf{H}(\mu)$

$$\sum_{n=1}^{\infty} \big|[\mathtt{j}(f_n)](h)\big|^2 = \sum_{n=1}^{\infty} \Big|\big\langle \overline{h}, \mathtt{j}(f_n)\big\rangle_{L^2(E,\mu)}\Big|^2 = \|\overline{h}\|^2_{L^2(E,\mu)} = \|h\|^2_{\mathbf{H}(\mu)} < \infty.$$

□

3.5 Recapitulation for Separable Frechet Spaces

We want to briefly recapitulate the theory up to this point in its most natural setting: separable Frechet spaces, i.e. Polish locally convex spaces.

Let (E,τ) be separable Frechet. The weak sigma-algebra $\sigma(\mathcal{C}(E)) = \sigma(E; E^*)$ coincides with the Borel-sigma-algebra and many other reasonable choices (Thm. 2.5). The mean and covariance operator of a measure on that sigma-algebra are continuous w.r.t. the Mackey-topology of the pairing $\langle E^*, E\rangle$ and can thus be represented as evaluations (Thm. 3.9). The characteristic functional of a finite signed measure completely characterizes that measure (Thm. 2.15), and the characteristic functional of a Gaussian measure can be given explicitly (Thm. 3.4). On separable Hilbert spaces, the Minlos–Sazanov Theorem (Thm. 2.14) gives an analogue of Bochner's Theorem (Thm. 2.13), but even on just Banach spaces, there is no such analogue. Due to Fernique's Theorem (Thm. 3.5), semi-norms are exponentially integrable w.r.t. a Gaussian measure and explicit tail bounds can be given. There exists a linear subspace $\mathbf{H}(\mu) \subseteq E$ s.t. for every $h \in \mathbf{H}(\mu)$ the measures μ and μ_h are equivalent (CM 1, Thm. 3.17). Furthermore, shifting along any element $x \notin \mathbf{H}(\mu)$ results in a mutually singular measure (CM 2, Thm. 3.19). The space $\big(\mathbf{H}(\mu), \|\cdot\|_{\mathbf{H}(\mu)}\big)$ with its associated norm is isometrically isomorphic to the reproducing kernel Hilbert space $\big(\mathbf{K}(\mu), \langle\cdot,\cdot\rangle_{L^2(E,\mu)}\big)$, which turns $\big(\mathbf{H}(\mu), \|\cdot\|_{\mathbf{H}(\mu)}\big)$ into a Hilbert space. The inclusion

$$\big(\mathbf{H}(\mu), \|\cdot\|_{\mathbf{H}(\mu)}\big) \hookrightarrow E$$

is weakly compact (Prop. 3.20) and $\mathbf{H}(\mu)$ is the intersection of all measurable linear subspaces of E with full measure, but $\mu[\mathbf{H}(\mu)] = 0$ unless $\mathbf{K}(\mu)$ has finite dimension (Prop. 3.21).

4 Abstract Wiener Space

In the preceding chapter we started with a locally convex TVS E and a Gaussian measure μ with covariance form $\mathfrak{q}_\mu$. Then we defined a Hilbert space embedded into $L^2(E,\mu)$ whose inner product was given by $\mathfrak{q}_\mu$ and subsequently identified an isometrically isomorphic subspace of E, the Cameron–Martin space $\mathbf{H}(\mu)$ with inner product $\langle\cdot,\cdot\rangle_{\mathbf{H}(\mu)}$. Heuristically speaking, the Gaussian measure μ on E was then given by

$$\frac{1}{\beta}\exp\left\{-\frac{1}{2}\|x\|^2_{\mathbf{H}(\mu)}\right\}\mathcal{D}x, \quad x \in E,$$

where $\mathcal{D}$ is the "infinite-dimensional Lebesgue measure on E" and β a normalization constant (see subsection 3.3.2 and 3.3.3). In this chapter we want to study the dual situation, which is the setting of the introductory section 1.2. We are given an expression that should be a measure μ with density w.r.t. $\mathcal{D}$, and, formally, looks like

$$\frac{1}{\beta}\exp\left\{-\frac{1}{2}\langle x,x\rangle_H\right\}\mathcal{D}x,$$

as well as a separable Hilbert space whose elements are precisely those for which $\langle x,x\rangle_H$ is finite. The goal is then to construct a TVS E on which this measure is well-defined, taking the role of $C_0[0,1]$ in the case of the classical Wiener space.

4.1 Classical Wiener Space

Before entering into a general discussion we first want to consider the classical example, the construction of which we will try to emulate. Recall that standard Brownian motion (BM) is defined as a stochastic process $B = (B_t)_{t\in[0,1]}$ on a probability space $(\Omega,\mathcal{A},\mathbb{P})$ s.t.

(a) $\mathbb{P}(B_0 = 0) = 1$,

(b) for any $1 \le i \le n$ and $0 = t_0 < t_1 < \ldots < t_n \le 1$ the random variables $B_{t_i} - B_{t_{i-1}}$ are independent with distribution $\mathcal{N}(0, t_i - t_{i-1})$, i.e. the increments of the process are centred, independent, and normally distributed with variance proportional to the length of the increment,

(c) $t \mapsto B_t(\omega)$ is continuous for $\mathbb{P}$-a.e. $\omega \in \Omega$.

Since for every $t \in [0,1]$ the mapping $\omega \mapsto B_t(\omega)$ is a random variable, the mapping $\omega \mapsto (B_t(\omega))_{t\in[0,1]}$ is measurable as a function

$$B : (\Omega,\mathcal{A},\mathbb{P}) \to \Big(\mathbb{R}^{[0,1]}, \sigma\big(\mathbb{R}^{[0,1]}, \{\mathrm{ev}_t\}_{t\in[0,1]}\big)\Big)$$

Furthermore, by the definition of BM, B takes values in the subspace $C_0[0,1]$ of continuous functions starting at 0 $\mathbb{P}$-a.s., which, equipped with the $\|\cdot\|_\infty$-norm, is a separable Banach space. For the same reason as above

$$B : (\Omega,\mathcal{A},\mathbb{P}) \to \Big(C_0[0,1], \sigma\big(C_0[0,1], \{\mathrm{ev}_t\}_{t\in[0,1]}\big)\Big)$$

is measurable, and since $\{\mathrm{ev}_t\}_{t\in[0,1]}$ separates points of $C_0[0,1]$, the weak sigma-algebra $\sigma(C_0[0,1], \{\mathrm{ev}_t\}_{t\in[0,1]})$ and the Borel sigma-algebra $\mathcal{B}(C_0[0,1])$ coincide by Theorem 2.5.

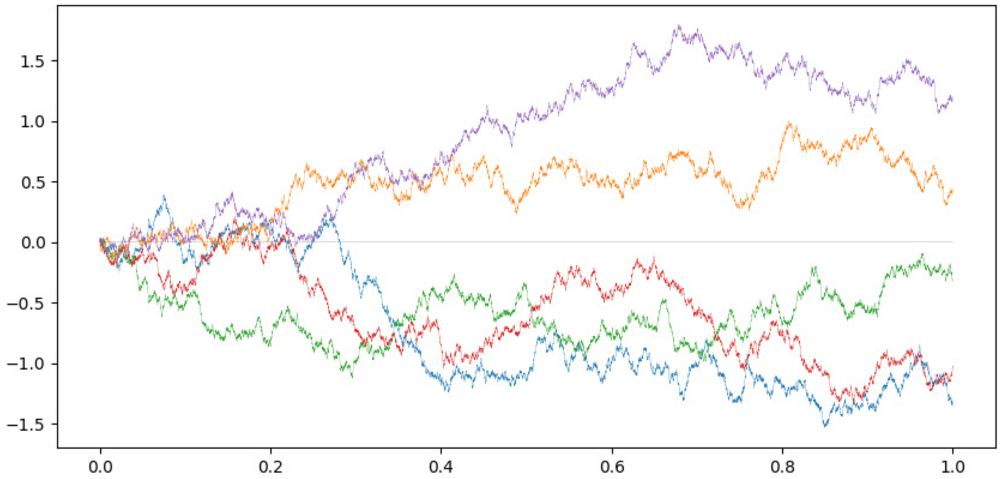

Figure 5: Sample paths of standard Brownian motion.

The distribution of B, a probability measure on $\mathcal{B}(C_0[0,1])$, is called the **Wiener measure** on $C_0[0,1]$. Note that any random variable with values in $C_0[0,1]$ whose law is the Wiener measure is necessarily a BM, i.e. a Brownian motion is a stochastic process with the Wiener measure as its distribution. A priori, it is not clear that such process (or equivalently such a measure) even exists in the way described, i.e. satisfying properties (a) - (c). In fact, the physical phenomenon of Brownian motion had been observed and described by R. Brown [3] almost 100 years prior to N. Wiener's proof of existence in the mathematical sense [32]. The proof of existence presented here will be constructive and serves as a prototype for the more general setting of abstract Wiener spaces.

Theorem 4.1. *Brownian Motion exists.*

Proof. The proof is taken from [20, Sec. 2.3.], the sleek title from R. Durrett. Let $I(n)$ denote the set of odd integers between 0 and 2^n with $n \in \mathbb{N}$. Let $\{\xi_{n,k} : k \in I(n), n \in \mathbb{N}\}$ be a sequence of iid $\mathcal{N}(0,1)$-distributed random variables defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We want to construct BM as a path-wise limit of sums of random variables

$$\sum_{n=1}^{N} \xi_{n,k}\ s_{n,k},$$

where $\{s_{n,k} : k \in I(n), n \in \mathbb{N}\}$ are the Schauder functions on $[0,1]$, defined by $s_{n,k}(t) := \int_0^t h_{n,k}(s)\,\mathrm{d}s$, $0 \le t \le 1$, and $\{h_{n,k} : k \in I(n), n \in \mathbb{N}\}$ are the Haar wavelets defined by

$$h_{n,k}(t) := \begin{cases} 2^{\frac{n-1}{2}}, & \frac{k-1}{2^n} \le t < \frac{k}{2^n} \\ -2^{\frac{n-1}{2}}, & \frac{k}{2^n} \le t < \frac{k+1}{2^n}\,, \\ 0, & \text{otherwise} \end{cases} \qquad k \in I(n), n \in \mathbb{N}$$

and $h_{0,0} \equiv 1$. Since the Haar wavelets form an ONB of $L^2[0,1]$, the Schauder functions form an ONB of the first Hilbert–Sobolev space $H_0^1[0,1]$. Define for every $N \in \mathbb{N}$ the function $B^{(N)} : (\Omega, \mathcal{A}, \mathbb{P}) \to C_0[0,1]$ via

$$B^{(N)}(\omega) = \sum_{n=1}^{N} \sum_{k \in I(n)} \xi_{n,k}(\omega) s_{n,k}, \quad B^{(0)}(\omega) = 0, \quad \omega \in \Omega.$$

$B^{(N)}(\omega)$ is continuous as a function of t, starts at 0, and the assignment $\omega \mapsto B^{(N)}(\omega)$ is weakly measurable and thus Borel measurable. We want to show that for $\mathbb{P}$-a.e. $\omega \in \Omega$ the sequence $(B^{(N)}(\omega))_{N\in\mathbb{N}}$ converges uniformly to a continuous function $\lim_{N\to\infty} B^{(N)}(\omega)$ s.t. the random variable $\omega \mapsto \lim_{N\to\infty} B^{(N)}(\omega)$ is a BM. Firstly, note that for any $n \in \mathbb{N}, k \in I(n)$ we have

$$\mathbb{P}\left[|\xi_{n,k}| > n\right] = \sqrt{\frac{2}{\pi}} \int_n^\infty e^{\frac{-u^2}{2}} \mathrm{d}u \le \sqrt{\frac{2}{\pi}} \int_n^\infty \frac{u}{n} e^{\frac{-u^2}{2}} \mathrm{d}u \le \sqrt{\frac{2}{\pi}} \frac{e^{\frac{-n^2}{2}}}{n} \quad ,$$

and thus

$$\mathbb{P}\left[\sup_{k\in I(n)} |\xi_{n,k}| > n\right] = \mathbb{P}\left[\bigcup_{k\in I(n)} \{|\xi_{n,k}| > n\}\right] \le \sum_{k\in I(n)} \sqrt{\frac{2}{\pi}} \frac{e^{\frac{-n^2}{2}}}{n} \le \sqrt{\frac{2}{\pi}} 2^n \frac{e^{\frac{-n^2}{2}}}{n}. \tag{4.1}$$

Since the last term in line (4.1) is summable, the Borel–Cantelli Lemma implies that

$$\mathbb{P}\left[\sup_{k\in I(n)} |\xi_{n,k}| > n \text{ for infinitely many } n\right] = 0$$

and hence

$$\mathbb{P}\left[\sup_{k\in I(n)} |\xi_{n,k}| \le n \text{ for all but finitely many } n\right] = 1.$$

In other words, for $\mathbb{P}$-a.e. $\omega \in \Omega$ there exists an $n_0(\omega) \in \mathbb{N}$ s.t. $\forall n \ge n_0(\omega)$: $\sup_{k\in I(n)} |\xi_{n,k}| \le n$. Thus for $\mathbb{P}$-a.e. $\omega \in \Omega$

$$\left\| \sum_{n=n_0(\omega)}^{\infty} \sum_{k\in I(n)} |\xi_{n,k}(\omega) s_{n,k}| \right\|_\infty \le \sum_{n=n_0(\omega)}^{\infty} \sum_{k\in I(n)} |\xi_{n,k}(\omega)| \|s_{n,k}\|_\infty \le \sum_{n=n_0(\omega)}^{\infty} n 2^{\frac{-(n+1)}{2}} < \infty,$$

where in the second to last estimate we used that

$$\|s_{n,k}\|_\infty = 2^{\frac{-(n+1)}{2}}, \quad \forall k \in I(n), n \in \mathbb{N}.$$

Hence the sequence $(B^{(N)})_{n\in\mathbb{N}}$ is absolutely convergent in the uniform norm $\mathbb{P}$-a.s.

To show that B is indeed a BM, we check the conditions of the definition. As a uniform limit of continuous functions starting at 0 the function $B(\omega)$ is also continuous and starts at 0 for $\mathbb{P}$-a.e. $\omega \in \Omega$, showing (a) and (c). To check condition (b), let $1 \le i \le n$ and $0 = t_0 < t_1 < \ldots < t_n \le 1$ be arbitrary. We will show that the characteristic function of the random vector $(B_{t_1} - B_{t_0}, \ldots, B_{t_n} - B_{t_{n-1}})$ coincides with that of a Gaussian random vector with independent and $\mathcal{N}(0, t_i - t_{i-1})$-distributed entries. That is, we want to show that for any $\alpha_1, \dots, \alpha_d \in \mathbb{R}$

$$\mathbb{E}\left[\exp\left\{i\sum_{i=1}^d \alpha_i(B_{t_i} - B_{t_{i-1}})\right\}\right] = \prod_{i=1}^d \exp\left\{-\frac{1}{2}\alpha_i^2(t_i - t_{i-1})\right\}.$$

Set $\alpha_{d+1} = 0$. Then rearranging the sum and using the Dominated Convergence Theorem leads to

$$\begin{aligned}
&\mathbb{E}\left[\exp\left\{i\sum_{i=1}^d \alpha_i(B_{t_i} - B_{t_{i-1}})\right\}\right] \\
=&\mathbb{E}\left[\exp\left\{-i\sum_{i=1}^d (\alpha_{i+1} - \alpha_i)B_{t_i}\right\}\right] \\
=&\lim_{N\to\infty}\mathbb{E}\left[\exp\left\{-i\sum_{i=1}^d (\alpha_{i+1} - \alpha_i)B_{t_i}^{(N)}\right\}\right] \\
=&\lim_{N\to\infty}\mathbb{E}\left[\exp\left\{-i\sum_{i=1}^d (\alpha_{i+1} - \alpha_i)\sum_{n=1}^N\sum_{k\in I(n)} \xi_{n,k}s_{n,k}(t_i)\right\}\right].
\end{aligned}$$

Independence of the $\xi_{n,k}$ and forming the characteristic functions further yield

$$\begin{aligned}
&= \lim_{N\to\infty}\prod_{n=1}^N\prod_{k\in I(n)}\mathbb{E}\left[\exp\left\{-i\xi_{n,k}\sum_{i=1}^d(\alpha_{i+1} - \alpha_i)s_{n,k}(t_i)\right\}\right] \\
&= \lim_{N\to\infty}\prod_{n=1}^N\prod_{k\in I(n)}\exp\left\{-\frac{1}{2}\left(\sum_{i=1}^d(\alpha_{i+1} - \alpha_i)s_{n,k}(t_i)\right)^2\right\} \\
&= \lim_{N\to\infty}\prod_{n=1}^N\prod_{k\in I(n)}\exp\left\{-\frac{1}{2}\left(\sum_{1\le i,j\le d}(\alpha_{i+1} - \alpha_i)(\alpha_{j+1} - \alpha_j)s_{n,k}(t_i)s_{n,k}(t_j)\right)\right\} \\
&= \lim_{N\to\infty}\exp\left\{-\frac{1}{2}\sum_{1\le i,j\le d}(\alpha_{i+1} - \alpha_i)(\alpha_{j+1} - \alpha_j)\sum_{n=1}^N\sum_{k\in I(n)}s_{n,k}(t_i)s_{n,k}(t_j)\right\} \\
&= \exp\left\{-\frac{1}{2}\sum_{0\le i,j\le d}(\alpha_{i+1} - \alpha_i)(\alpha_{j+1} - \alpha_j)\sum_{n=1}^\infty\sum_{k\in I(n)}s_{n,k}(t_i)s_{n,k}(t_j)\right\}.
\end{aligned}$$

Recall Parseval's identity: For f, g in a separable Hilbert space H and $\{e_n\}_{n\in\mathbb{N}}$ an ONB of H we have

$$\langle f, g\rangle = \sum_{n\in\mathbb{N}}\langle f, e_n\rangle_H\langle g, e_n\rangle_H.$$

Applied to the ONB of $L^2[0,1]$ consisting of the Haar wavelets and the indicator functions $f = 1_{[0,t_i]}, g = 1_{[0,t_j]}$ this yields

$$\sum_{n=1}^{\infty}\sum_{k\in I(n)} s_{n,k}(t_i)s_{n,k}(t_j) = \sum_{n=1}^{\infty}\sum_{k\in I(n)} \langle 1_{[0,t_i]}, h_{n,k}\rangle_{L^2}\langle 1_{[0,t_j]}, h_{n,k}\rangle_{L^2} = \langle 1_{[0,t_i]}, 1_{[0,t_j]}\rangle_{L^2} = t_i \wedge t_j.$$

Using this and rearranging the sums yields

$$\begin{aligned}
&= \exp\left\{-\frac{1}{2}\left(2\sum_{0\le i\neq j\le d}(\alpha_{i+1}-\alpha_i)(\alpha_{j+1}-\alpha_j)t_i\wedge t_j + \sum_{i=1}^{d}(\alpha_{i+1}-\alpha_i)^2 t_i\right)\right\}\\
&= \exp\left\{-\sum_{i=1}^{d-1}\sum_{j=i+1}^{d}(\alpha_{i+1}-\alpha_i)(\alpha_{j+1}-\alpha_j)t_i - \frac{1}{2}\sum_{j=1}^{d}(\alpha_{j+1}-\alpha_j)^2 t_j\right\}\\
&= \exp\left\{-\sum_{i=1}^{d-1}(\alpha_{i+1}-\alpha_i)t_i\underbrace{\sum_{j=i+1}^{d}(\alpha_{j+1}-\alpha_j)}_{=\alpha_{d+1}-\alpha_{j+1}=-\alpha_{j+1}} - \frac{1}{2}\sum_{j=1}^{d}(\alpha_{j+1}-\alpha_j)^2 t_j\right\}\\
&= \exp\left\{-\frac{1}{2}\sum_{i=1}^{d-1}(\alpha_{i+1}^2-\alpha_i^2)t_i - \frac{1}{2}\alpha_d^2 t_d\right\}\\
&= \exp\left\{\sum_{i=1}^{d} -\frac{1}{2}\alpha_i^2(t_i-t_{i-1})\right\} = \prod_{i=1}^{d}\exp\left\{-\frac{1}{2}\alpha_i^2(t_i-t_{i-1})\right\},
\end{aligned}$$

which is what was to be shown. □

4.2 Cylinder Measures

Back to the general case. The first step in the program is to define a "measure" on H, which controls the finite-dimensional distributions, and from which the actual measure on E is built. We write "measure" in quotation marks here, since it will only be defined on the algebra of cylinder sets of H (which is generally not a sigma-algebra) and will not extend to a bona fide measure on $\sigma(\mathcal{C}(H))$ (see Proposition 4.9).

Lemma 4.2. *Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Then cylinder sets can be assumed to be given by orthonormal functionals, i.e.*

$$\mathcal{C}(H) = \left\{C_{A,e_1,\dots,e_n} : n\in\mathbb{N}, A\in\mathcal{B}(\mathbb{R}^n), e_1,\dots,e_n\in H \text{ orthonormal}\right\}, \tag{4.2}$$

where $\mathcal{C}(H)$ is the algebra of cylinder sets of H, defined in Definition 2.1 and Proposition 2.2.

Proof. "$\supseteq$" clear.
"$\subseteq$" Let $C\in\mathcal{C}(H)$. Then there exist $x_1,\dots,x_m\in H$ and $A\in\mathcal{B}(\mathbb{R}^n)$ s.t.

$$C = \left\{h\in H : \big(\langle h,x_1\rangle,\dots,\langle h,x_m\rangle\big)\in A\right\}.$$

Let $K := (x_1, \dots, x_m)_H \subseteq H$ denote the subspace of H spanned by $x_1, \dots, x_m$ and choose an orthonormal basis $\{e_1, \dots, e_n\}$ of K where $\dim K = n$. Define a matrix $T \in \mathbb{R}^{m\times n}$ by

$$T_{ij} := \langle x_i, e_j \rangle, \quad 1 \le i \le m, 1 \le j \le n.$$

Then

$$\begin{aligned}
C &= \left\{ h \in H : \left(\left\langle h, \sum_{i=1}^{n} \langle x_1, e_j \rangle e_j \right\rangle, \dots, \left\langle h, \sum_{i=1}^{n} \langle x_m, e_j \rangle e_j \right\rangle \right) \in A \right\} \\
&= \left\{ h \in H : \left(\sum_{i=1}^{n} \langle x_1, e_j \rangle \langle h, e_j \rangle, \dots, \sum_{i=1}^{n} \langle x_m, e_j \rangle \langle h, e_j \rangle \right) \in A \right\} \\
&= \Big\{ h \in H : T\big(\langle h, e_1 \rangle, \dots, \langle h, e_n \rangle\big) \in A \Big\} \\
&= \Big\{ h \in H : \big(\langle h, e_1 \rangle, \dots, \langle h, e_n \rangle\big) \in T^{-1}(A) \Big\},
\end{aligned}$$

where T^{-1} denotes the pre-image under the map T. Since T is continuous, $T^{-1}(A)$ lies in $\mathcal{B}(\mathbb{R}^n)$ and thus C lies in the right hand side of formula (4.2). $\square$

Definition 4.3. Let (E, τ) be a locally convex TVS and $\mathcal{C}(E)$ its algebra of cylinder sets. A (probability) **cylinder measure** ν on E is a finitely additive set function $\mathcal{C}(E) \to [0, 1]$ s.t.

(i) $\nu(\emptyset) = 0$, $\nu(E) = 1$, and
(ii) for any continuous linear operator $P : (E, \tau) \to \mathbb{R}^n$ the push-forward of ν along P is a measure on $\mathcal{B}(\mathbb{R}^n)$ in the usual sense.

The Fourier transform of a cylinder measure is defined by

$$\widehat{\nu}(f) := \int_{\mathbb{R}} \exp\{it\} \mathrm{d}\big[\nu \circ f^{-1}\big](t), \quad f \in E^*.$$

Note that any probability measure μ defined on $\sigma(E; E^*)$ restricts to a cylinder measure $\mu|_{\mathcal{C}(E)}$ on $\mathcal{C}(E)$, and that $\widehat{\mu}$ and $\widehat{\mu|_{\mathcal{C}(E)}}$ agree on E^*: Let $f \in E^*$ be arbitrary. Then

$$\begin{aligned}
\widehat{\mu|_{\mathcal{C}(E)}}(f) &= \int_{\mathbb{R}} \exp\big\{it\big\} \,\mathrm{d}\Big[\underbrace{\mu|_{\mathcal{C}(E)} \circ f^{-1}}_{=\mu\circ f^{-1}}\Big](t) \\
&= \int_{\mathbb{R}} \exp\big\{it\big\} \,\mathrm{d}\left[\mu \circ f^{-1}\right](t) \\
&= \int_{E} \exp\big\{if(x)\big\} \,\mathrm{d}\mu(t) = \widehat{\mu}(f).
\end{aligned}$$

Definition 4.4. Let $(H, \langle \cdot, \cdot \rangle_H)$ be a separable Hilbert space and denote the standard Gaussian distribution $\mathcal{N}(0, 1^{n\times n})$ on $\mathbb{R}^n$ by γ_n. Define $\nu_H : \mathcal{C}(H) \to [0, 1]$ via

$$\nu_H(C) := \gamma_n(A), \quad A \in \mathcal{B}(\mathbb{R}^n)$$

where $C = P^{-1}(A)$ and $P : H \to \mathbb{R}^n$ is a continuous linear operator given by

$$P = \Big(\langle\cdot, e_1\rangle, \ldots, \langle\cdot, e_n\rangle\Big)$$

for an orthonormal set $\{e_i\}_{i=1}^n$. The function ν_H is called the **canonical (or standard) cylinder measure on** H.

Remark 4.5. We want to stress the fact that despite its name, the canonical cylinder measure ν_H on a separable Hilbert space H is not a measure, since it is only defined on an algebra of sets $\mathcal{C}(H)$, and, as will be shown in Proposition 4.9, does not extend to a bona fide measure on the generated sigma-algebra $\sigma\big(\mathcal{C}(H)\big)$.

From the definition alone, it is not clear that ν_H is a cylinder measure or even merely well-defined. This will be resolved in the next proposition.

Proposition 4.6. *Let $(H, \langle\cdot,\cdot\rangle_H)$ be a separable Hilbert space. Then its canonical cylinder measure ν_H is a well-defined cylinder measure and has Fourier transform*

$$\widehat{\nu_H}(x) = \exp\left\{-\frac{1}{2}\langle x, x\rangle_H\right\}.$$

Proof. **Well-definedness**, adapted from [8, Prop. 4.37.]: Lemma 4.2 shows that ν_H is defined for every $C \in \mathcal{C}(H)$. We want to show that $\nu_H(C)$ is independent of the choice of basis $A \in \mathbb{R}^n$ and generators $e_1, \ldots, e_n \in H$ of C. So assume

$$\Big\{h \in H : \big(\langle h, e_1\rangle, \ldots, \langle h, e_n\rangle\big) \in A\Big\} = C = \Big\{h \in H : \big(\langle h, f_1\rangle, \ldots, \langle h, f_m\rangle\big) \in B\Big\}$$

for some $e_1, \ldots, e_n \in H$ orthonormal, $f_1, \ldots, f_m \in H$ orthonormal, $A \in \mathbb{R}^n$, and $B \in \mathbb{R}^m$. Here we can assume orthonormality because of Lemma 4.2. Since the standard Gaussian distribution is a product measure we have

$$\gamma_n(A) = \gamma_{n+m}(A \times \mathbb{R}^m) \tag{4.3}$$
$$\gamma_m(B) = \gamma_{n+m}(B \times \mathbb{R}^n). \tag{4.4}$$

Hence if we are able to show that the two right hand sides above coincide, then the proposition is proven. To show this, complete $e_1, \ldots, e_n$ and $f_1, \ldots, f_m$ to orthonormal bases $e_1, \ldots, e_n, \ldots, e_{n+m}$ and $f_1, \ldots, f_m, \ldots, f_{n+m}$, respectively, of the subspace of H generated by $e_1, \ldots, e_n, f_1, \ldots, f_m$. Denote by T the matrix given by

$$T_{ij} := \langle e_i, f_j\rangle, \quad 1 \le i, j \le n+m.$$

The matrix is orthogonal since $e_1, \ldots, e_n, \ldots, e_{n+m}$ and $f_1, \ldots, f_m, \ldots, f_{n+m}$ are orthonormal bases. Denote by $P : H \to \mathbb{R}^{n+m}$ the surjective linear operator $\big(\langle\cdot, f_1\rangle, \ldots, \langle\cdot, f_{n+m}\rangle\big)$. Then we have

$$P^{-1}(B \times \mathbb{R}^n) = C = P^{-1}\big(T^{-1}(A \times \mathbb{R}^m)\big).$$

Since P is surjective we have $B \times \mathbb{R}^n = T^{-1}(A \times \mathbb{R}^m)$. Thus, since the standard Gaussian distribution is invariant under orthogonal transformations, the right hand side of equations (4.3) and (4.4) coincide and the set function ν_H is well-defined.

Cylinder measure, adapted from [8, Prop. 4.37.]: To show that the set function ν_H is a cylinder measure we check the axioms. We have

- $\nu_H(\emptyset) = \nu_H(P^{-1}(\emptyset)) = \gamma_n(\emptyset) = 0$ and $\nu_H(H) = \nu_H(P^{-1}(\mathbb{R}^n)) = \gamma_n(\mathbb{R}^n) = 1$ for some surjective bounded linear operator $P : H \to \mathbb{R}^n$.
- The push-forward along any bounded linear operator $Q : H \to \mathbb{R}^n$ is a measure, since the push-forward along any $P = \big(\langle \cdot, e_1\rangle, \dots, \langle \cdot, e_n\rangle\big)$ is a measure, namely γ_n.
- To show finite additivity let $C_1, \dots, C_m \in \mathcal{C}(H)$ be disjoint. By definition, for every $1 \le i \le m$ there exist a bounded, finite-dimensional linear operator $P_i : H \to \mathbb{R}^{n(i)}$ defined by

$$P_i = \Big(\langle \cdot, e_1^i\rangle, \dots, \langle \cdot, e_{n(i)}^i\rangle\Big), \quad e_1^i, \dots, e_{n(i)}^i \in H, \text{orthonormal},$$

and a Borel set $A_i \in \mathcal{B}(\mathbb{R}^{n(i)})$ s.t. $C_i = P_i^{-1}(A_i)$. We want to find a single $\overline{n} \in \mathbb{N}$ and $\overline{P} : H \to \mathbb{R}^{\overline{n}}$ s.t. $C_i = \overline{P}^{-1}(A_i')$ for some (possibly different) A_i', $1 \le i \le m$. For this, we pursue a similar strategy as for the well-definedness. For every $1 \le i \le m$ complete $e_1^i, \dots, e_{n(i)}^i \in H$ to an ONB of the span of $K := \{e_1^1, \dots, e_{n(1)}^1, \dots, e_1^m, \dots, e_{n(m)}^m\}$ and define by T_i the orthogonal linear operator $\mathbb{R}^{\overline{n}} \to \mathbb{R}^{\overline{n}}$ transforming the first completed basis into the i-th completed basis. Then we have

$$\begin{aligned}
C_i &= P_i^{-1}(A_i) \\
&= \Big\{x \in H : \big(\langle x, e_1^i\rangle, \dots, \langle x, e_{n(i)}^i\rangle\big) \in A_i\Big\} \\
&= \Big\{x \in H : \big(\langle x, e_1^i\rangle, \dots, \langle x, e_{n(i)}^i\rangle, \dots, \langle x, e_{\overline{n}}^i\rangle\big) \in A_i \times \mathbb{R}^{\overline{n}-n(i)}\Big\} \\
&= \Big\{x \in H : \big(\langle x, e_1^1\rangle, \dots, \langle x, e_{n(1)}^1\rangle, \dots, \langle x, e_{\overline{n}}^1\rangle\big) \in \underbrace{T_i^{-1}(A_i \times \mathbb{R}^{\overline{n}-n(i)})}_{=:A_i'}\Big\}
\end{aligned} \tag{4.5}$$

and define

$$\overline{P} := \Big(\langle \cdot, e_1^1\rangle, \dots, \langle \cdot, e_{n(1)}^1\rangle, \dots, \langle \cdot, e_{\overline{n}}^1\rangle\Big).$$

Thus equation (4.5) becomes $C_i = \overline{P}^{-1}(A_1')$. Note that since the C_i are disjoint, so are the A_i' for $1 \le i \le m$. With this in place, we deduce

$$\begin{aligned}
\nu_H\left(\bigcup_{i=1}^m C_i\right) &= \nu_H\left(\bigcup_{i=1}^n \overline{P}^{-1}(A_i')\right) = \nu_H\left(\overline{P}^{-1}\bigcup_{i=1}^m A_i'\right) = \gamma_{\overline{n}}\left(\bigcup_{i=1}^m A_i'\right) \\
&= \sum_{i=1}^m \gamma_{\overline{n}}(A_i') = \sum_{i=1}^m \gamma_{n(i)}(A_i) = \sum_{i=1}^m \nu_H(C_i)
\end{aligned}$$

where in the second to last equality we used the invariance of the standard Gaussian measure under orthogonal transformations and the fact that the standard Gaussian measure is a product measure, i.e.

$$\gamma_{\overline{n}}\big(A_i'\big) = \gamma_{\overline{n}}\Big(T_i^{-1}\big(A_i \times \mathbb{R}^{\overline{n}-n(i)}\big)\Big) = \gamma_{\overline{n}}\Big(A_i \times \mathbb{R}^{\overline{n}-n(i)}\Big) = \gamma_{n(i)}\big(A_i\big).$$

This shows the finite additivity. Note that the proof relies on the existence of $\overline{n}$ as above and thus cannot be adapted to show countable additivity.

Fourier transform: Let $x \in H$ be arbitrary. Then, since $\langle \cdot, x\rangle : H \to \mathbb{R}$ is a Gaussian random variable with variance $\operatorname{Var}(x) = \|x\|^2$, $\widetilde{\nu_H}\big(\langle \cdot, x\rangle\big)$ equals

$$\int_{\mathbb{R}} \exp\{it\}\mathrm{d}\left[\nu_H \circ \langle \cdot, x\rangle^{-1}\right](t) = \phi_{\langle \cdot, x\rangle}(1) = \exp\left\{-\frac{1}{2}\operatorname{Var}(x)\right\} = \exp\left\{-\frac{1}{2}\langle x, x\rangle\right\}.$$

□

Remark 4.7. The standard Gaussian distribution on $\mathbb{R}^n$ is the canonical cylinder measure of $(\mathbb{R}^n, \langle \cdot, \cdot\rangle_{\mathbb{R}^n})$. The Gaussian measure $\mathcal{N}(0, \Sigma)$ is the canonical cylinder measure of $(\mathbb{R}^n, \langle \cdot, \Sigma \cdot\rangle_{\mathbb{R}^n})$.

Lemma 4.8. *Let $(H, \langle \cdot, \cdot\rangle)$ be a separable Hilbert space, $\{e_n\}_{n\in\mathbb{N}}$ an ONB of H and ν a cylinder measure on $\mathcal{C}(H)$. Then the following are equivalent:*

(a) ν is the canonical cylinder measure ν_H.

(b) For any finite subset $\{e_{i_j}\}_{j=1}^m$ of $\{e_n\}_{n\in\mathbb{N}}$ the bounded linear operator $\big(\langle \cdot, e_{i_1}\rangle, \dots, \langle \cdot, e_{i_m}\rangle\big)$ is Gaussian with distribution $\mathcal{N}(0, 1^{n\times n})$.

(c) For any $x \in H$ the functional $\langle \cdot, x\rangle$ is Gaussian with distribution $\mathcal{N}(0, \|x\|_H^2)$.

(d) For any $x \in H$ with norm 1 the functional $\langle \cdot, x\rangle$ is standard Gaussian, i.e. has distribution $\mathcal{N}(0, 1)$.

Proof. "(a) ⇒ (b) ⇒ (c) ⇒ (d)" By definition.
"(b) ⇒ (a)" Let $\{f_1, \dots, f_m\} \subseteq H$ be an orthonormal set. Then the random vector $\big(\langle \cdot, f_1\rangle, \dots, \langle \cdot, f_m\rangle\big)$ equals

$$\begin{pmatrix} \left\langle \cdot, \sum_{n=1}^{\infty}\langle f_1, e_n\rangle e_n \right\rangle \\ \vdots \\ \left\langle \cdot, \sum_{n=1}^{\infty}\langle f_m, e_n\rangle e_n \right\rangle \end{pmatrix} = \lim_{N\to\infty} \begin{pmatrix} \left\langle \cdot, \sum_{n=1}^{N}\langle f_1, e_n\rangle e_n \right\rangle \\ \vdots \\ \left\langle \cdot, \sum_{n=1}^{N}\langle f_m, e_n\rangle e_n \right\rangle \end{pmatrix} = \lim_{N\to\infty} A^{(N)} \underbrace{\begin{pmatrix} \langle \cdot, e_1\rangle \\ \vdots \\ \langle \cdot, e_N\rangle \end{pmatrix}}_{\sim\mathcal{N}(0, 1^{N\times N})}$$

where $A^{(N)} \in \mathbb{R}^{m\times N}$. By assumption the vector on the right hand side has distribution $\mathcal{N}(0, 1^{N\times N})$ and hence $\big(\langle \cdot, f_1\rangle, \dots, \langle \cdot, f_m\rangle\big)$ is the norm-limit of Gaussian random variables with distribution $\mathcal{N}(0, A^{(N)}A^{(N)^T})$. Because norm convergence implies weak* convergence (and thus a.s. convergence), this shows that $\big(\langle\cdot,f_1\rangle,\dots,\langle\cdot,f_m\rangle\big)$ is Gaussian with distribution $\mathcal{N}(0,\lim_{N\to\infty}A^{(N)}A^{(N)^T})$.[10] For any $N\in\mathbb{N}$ and $1\le i,j\le m$ we have

$$\lim_{N\to\infty}\left(A^{(N)}A^{(N)^T}\right)_{i,j}=\lim_{N\to\infty}\left\langle\begin{pmatrix}\langle f_i,e_1\rangle\\ \vdots\\ \langle f_i,e_N\rangle\end{pmatrix},\begin{pmatrix}\langle f_j,e_1\rangle\\ \vdots\\ \langle f_j,e_N\rangle\end{pmatrix}\right\rangle=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}, \tag{4.6}$$

where the last line is due to the linear isometry $H\to\ell^2$ via $x\mapsto(\langle x,e_n\rangle)_{n\in\mathbb{N}}$ and the fact that f_i and f_j are orthogonal w.r.t. the inner product in H. Thus $\lim_{N\to\infty}A^{(N)}A^{(N)^T}$ is the identity matrix $1^{n\times n}$, which implies that $\langle\cdot,f_1\rangle,\dots,\langle\cdot,f_m\rangle$ are independent.

"(c) $\Rightarrow$ (b)" Assume (c) was true, but (b) was not. Note that $\big(\langle\cdot,e_{i_1}\rangle,\dots,\langle\cdot,e_{i_m}\rangle\big)$ is a Gaussian vector since for any $(\alpha_i)_{i=1}^m\in\mathbb{R}^m$ we have

$$\left\langle\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_m\end{pmatrix},\begin{pmatrix}\langle\cdot,e_{i_1}\rangle\\ \vdots\\ \langle\cdot,e_{i_m}\rangle\end{pmatrix}\right\rangle=\sum_{k=1}^m\alpha_i\langle\cdot,e_{i_k}\rangle=\Big\langle\cdot,\underbrace{\sum_{k=1}^m\alpha_k e_{i_k}}_{\in H}\Big\rangle$$

which is Gaussian by assumption. Since (b) is not true by assumption, there exist $e_1,e_2\in H$ orthonormal s.t. $\mathrm{Cov}(e_1,e_2)\neq0$, and thus

$$\begin{aligned}\mathrm{Var}\left(\frac{1}{\sqrt2}(e_1+e_2)\right)&=\int_H\left\langle x,\frac{1}{\sqrt2}(e_1+e_2)\right\rangle^2\mathrm{d}\nu(x)\\&=\frac12\int_H\Big(\langle x,e_1\rangle+\langle x,e_2\rangle\Big)^2\mathrm{d}\nu(x)\\&=\frac12\Big(\underbrace{\int_H\langle x,e_1\rangle^2\mathrm{d}\nu(x)}_{=\mathrm{Var}(e_1)=1}+2\underbrace{\int_H\langle x,e_1\rangle\langle x,e_2\rangle\mathrm{d}\nu(x)}_{=\mathrm{Cov}(e_1,e_2)\neq0}+\underbrace{\int_H\langle x,e_2\rangle^2\mathrm{d}\nu(x)}_{=\mathrm{Var}(e_1)=1}\Big)\\&=1+\mathrm{Cov}(e_1,e_2)\neq1,\end{aligned}$$

despite the fact that

$$\left\|\frac{1}{\sqrt2}(e_1+e_2)\right\|^2=\frac12\langle e_1+e_2,e_1+e_2\rangle=\frac12\Big(\underbrace{\|e_1\|_H^2}_{=1}+2\underbrace{\langle e_1,e_2\rangle}_{=0}+\underbrace{\|e_2\|_H^2}_{=1}\Big)=1,$$

which is a contradiction to (c).

"(d) $\Rightarrow$ (c)" Clear.

□

Proposition 4.9. *Let $(H,\langle\cdot,\cdot\rangle_H)$ be a separable Hilbert space. If H is infinite-dimensional, then its canonical cylinder measure ν_H does not extend to a σ-additive measure on $\sigma(\mathcal{C}(H))$.*

[10] To be precise, this follows only after we show that the limit $\lim_{N\to\infty}A^{(N)}A^{(N)^T}$ exists, which is done in equation (4.6).

We give two arguments:

Proof. (1) This follows by Proposition 3.10. Assume ν_H did extend to a σ-additive measure $\overline{\nu_H}$ on $\sigma(\mathcal{C}(H))$. Then the Fourier transform $\widehat{\nu_H}$ of ν_H and the characteristic functional $\widehat{\overline{\nu_H}}$ of $\overline{\nu_H}$ would coincide. Hence $\overline{\nu_H}$ would be a measure on $\sigma(\mathcal{C}(H))$ with non-compact covariance operator, which cannot exist by Proposition 3.10.

□

Proof. (2), [8, Prop. 4.38.]: Let $\{e_n\}_{n\in\mathbb{N}}$ be an ONB of H and define the cylinder sets

$$A_{n,k} := \Big\{x \in H : |\langle x, e_i\rangle| \le k, 0 \le i \le n\Big\} = \begin{pmatrix} \langle \cdot, e_1\rangle \\ \vdots \\ \langle \cdot, e_n\rangle \end{pmatrix}^{-1} [-k,k]^n, \quad 0 \le n, k.$$

Note that $B_k(0) \subseteq A_{n,k}$ for every $n \in \mathbb{N}$ and that

$$\nu_H(A_{n,k}) = \gamma_n\big([-k,k]^n\big) = \gamma\big([-k,k]\big)^n$$

where the last equality is due to the fact that γ_n is a product measure. Since $\gamma\big([-k,k]\big) < 1$, for any $k \in \mathbb{N}$ we may choose an $n_k \in \mathbb{N}$ large enough s.t.

$$\nu_H(A_{n_k,k}) = \gamma\big([-k,k]\big)^{n_k} < 2^{-k},$$

thus

$$\sum_{n=1}^{\infty} \nu_H(A_{n_k,k}) < \sum_{n=1}^{\infty} 2^{-k} = 1,$$

and hence

$$H = \bigcup_{k=1}^{\infty} B_k(0) \subseteq \bigcup_{k=1}^{\infty} A_{n_k,k} \subseteq H.$$

Therefore, if ν_H had a σ-additive extension to $\sigma(\mathcal{C}(H))$, then

$$1 = \nu_H(H) = \nu_H\left(\bigcup_{k=1}^{\infty} A_{n_k,k}\right) \le \sum_{n=1}^{\infty} \nu_H(A_{n_k,k}) < 1,$$

which is a contradiction. □

The latter of the two proofs illustrates the problem: a product probability measure s.t. μ has positive mass outside of $[-k,k]$ will concentrate outside of $[-k,k]^n$ for large n. In other words, the mass of the supposed μ is concentrated "at infinity". We thus continue our program by trying to interpret H as a subspace of some larger space E that includes the entire mass of μ. We want to illustrate this via the following example.

Example 4.10 ($(X_n)_{n\in\mathbb{N}}$ on a weighted sequence space)**.** Consider the separable Hilbert space $(\ell^2(\mathbb{R}), \|\cdot\|_2)$ of real-valued, square-summable sequences and let ν_{ℓ^2} denote its canonical cylinder measure. By Lemma 4.8 this means that for any $n \in \mathbb{N}$ the vector $(\mathrm{ev}_{i_1}, \ldots, \mathrm{ev}_{i_n})$ is standard Gaussian. One should think of ν_{ℓ^2} as the distribution of a sequence $(X_n)_{n\in\mathbb{N}}$ of iid $\mathcal{N}(0,1)$-distributed random variables on some probability space $(\Omega,\mathcal{A},\mathbb{P})$. Strictly speaking, $(X_n)_{n\in\mathbb{N}}:\Omega\to\ell^2$ is not well-defined, since $\sum_{n=1}^\infty X_n^2=\infty$ $\mathbb{P}$-a.s. So in accordance with Proposition 4.9 we have $\nu_{\ell^2}(\ell^2)=0$. Generic elements of $(X_n)_{n\in\mathbb{N}}$, i.e. those that carry the weight of ν_{ℓ^2}, have infinite $\|\cdot\|_2$-norm. So our strategy is to consider a new norm for which generic elements of $(X_n)_{n\in\mathbb{N}}$ have finite norm, that is, a norm that "detects the mass at infinity".

Consider the sequence $a:=(a_n)_{n\in\mathbb{N}}=(2^{-n})_{n\in\mathbb{N}}\in\ell^2$ and define ℓ^2_a as the a-weighted ℓ^2 space, i.e.

$$\ell^2_a:=\left\{(x_n)_{n\in\mathbb{N}}\in\mathbb{R}^{\mathbb{N}}:\sum_{n=1}^\infty 2^{-n}x_n^2<\infty\right\}$$

together with the inner product given by $\langle x,y\rangle_a=\sum_{n=1}^\infty 2^{-n}x_ny_n$. Then, formally,

$$\nu_{\ell^2}\left(\ell^2_a\right)=\mathbb{P}\Big((X_n)_{n\in\mathbb{N}}\in\ell^2_a\Big)=\mathbb{P}\left(\sum_{n=1}^\infty 2^{-n}X_n^2<\infty\right)=1. \tag{4.7}$$

But how is the first expression in formula (4.7) even defined, given that ν_{ℓ^2} is a cylinder measure on ℓ^2, not ℓ^2_a? And how (or rather in what sense) does this extend ν_{ℓ^2} as defined on $\mathcal{C}(\ell^2)$? These questions will be answered with the general construction of abstract Wiener spaces due to L. Gross [15].

4.3 Construction via Measurable Norms

Recall the strategy alluded to in the preceding section. We want to construct a measure on H with the property that orthonormal elements are standard Gaussian. We cannot have that though, since the mass of such a measure would be concentrated at infinity. The solution is to embed H into a larger space E which includes those places at infinity, i.e. E is the closure of H in a suitable sense.

Recall Example 4.10 and the heuristic that the canonical cylinder measure there was just "the distribution" of a sequence of iid standard normal random variables. More precisely, we considered the ONB $(\delta_n)_{n\in\mathbb{N}}$ of ℓ^2, where δ_n is the sequence having 1 in the n-th position and 0 in the others, and ν_{ℓ^2} as the supposed distribution of the random variable

$$Z:=\sum_{k=1}^\infty X_n\delta_n=\big(X_1,X_2,\dots\big).$$

However, in the $\|\cdot\|_2$-norm this series does not converge $\mathbb{P}$-a.s., while in the $\|\cdot\|_a$-norm it does. We want to replicate this process more abstractly. We will consider

$$Z:=\sum_{n=1}^\infty \xi_ne_n,$$

where $\{\xi_n\}_{n\in\mathbb{N}}$ is a sequence of iid $\mathcal{N}(0,1)$ real-valued random variables and $\{e_n\}_{n\in\mathbb{N}}$ is an ONB of H, and find a suitable norm (analogous to $\|\cdot\|_a$) s.t. the series converges $\mathbb{P}$-a.s.

Definition 4.11. Let $(H, \langle\cdot,\cdot\rangle_H)$ be a separable Hilbert space, ν_H its canonical cylinder measure, and $\mathcal{P}(H)$ the set of finite-dimensional, orthogonal projections on H. A norm $\|\cdot\|$ on H is called **measurable** w.r.t. (H, ν_H) (in the sense of Gross) if for every $\varepsilon > 0$ there exists a finite-dimensional orthogonal projection operator $P_\varepsilon \in \mathcal{P}(H)$ s.t. for any $P \in \mathcal{P}(H)$ with $P \perp P_\varepsilon$ we have

$$\nu_H\Big\{x \in H : \|Px\| > \varepsilon\Big\} < \varepsilon.$$

Remark 4.12. ([8, p. 27]) The name "measurable norm" is perhaps a bit misleading. It does not say anything about $\|\cdot\|$ being measurable as a function $H \to \mathbb{R}$, but rather it expresses that $\|\cdot\|$ is compatible with μ in some sense.

Proposition 4.13. *Let $(H_0^1[0,1], \langle\cdot,\cdot\rangle_{H_0^1[0,1]})$ be the first Hilbert–Sobolev space. Then $\|\cdot\|_\infty$ is a measurable norm w.r.t. $(H_0^1[0,1], \nu_{H_0^1[0,1]})$, while $\|\cdot\|_{H_0^1[0,1]}$ is not.*

Proof. See [8, Lem. 4.41., Prop. 4.42.]. □

Theorem 4.14. *Let $(H, \langle\cdot,\cdot\rangle_H)$ be a separable Hilbert space, ν_H its canonical cylinder measure, $\|\cdot\|_E$ a measurable norm w.r.t. (H, ν_H), E the completion of H under the $\|\cdot\|_E$-norm and $\boldsymbol{i} : H \hookrightarrow E$ the inclusion of H into E. Then there exists a measure μ on $\mathcal{B}(E)$ s.t.*

(1) μ is an extension of $\nu_{H*} := \nu_H \circ \boldsymbol{i}^{-1}$, which is defined on $\mathcal{C}(E) \subseteq \mathcal{B}(E)$

(2) $(H, \|\cdot\|_H)$ coincides with the Cameron–Martin space of (E, μ).

Proof. (1): We define μ as the distribution of a random variable Z, which will be constructed as in Theorem 4.1. Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of iid, real-valued, $\mathcal{N}(0,1)$-distributed random variables, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Inductively define a sequence of orthogonal projections as follows.

Define $P_0 = 0$ and assume P_n was defined. Then, since $\|\cdot\|_E$ is a measurable norm, there exists a finite rank orthogonal projection operator P_n' s.t.

$$\mu\big\{x \in H : \|Px\|_E > 2^{-n}\big\} < 2^{-n}, \quad \forall P \in \mathcal{P}(H) \ \text{ s.t. } \ P \perp P_n'. \tag{4.8}$$

Now define

$$P_n = P_{n-1} + P_n' + \langle e_n', \cdot\rangle_H$$

where e_n lies in the orthogonal complement of $\big(\sum_{i=1}^{n-1} P_i\big)(H)$. From

$$P_n' \subseteq P_{n-1} + P_n' + \langle e_n', \cdot\rangle_H$$

we deduce that any $P \in \mathcal{P}(H)$ which is orthogonal to $P_{n-1}+P_n'+\langle e_n', \cdot\rangle_H$ is also orthogonal to P_n'. Hence any $P \in \mathcal{P}(H)$ s.t. $P \perp \big(P_{n-1} + P_n' + \langle e_n', \cdot\rangle_H\big)$ also fulfils the condition in formula (4.8). Thus the sequence P_n is increasing and for each $n \in \mathbb{N}$ the orthogonal projection P_n fulfils the condition in formula (4.8). By adding $\langle e_n', \cdot\rangle_H$ we ensured that $P_n \uparrow I$.[11] Now choose an ONB $(e_k)_{k\in\mathbb{N}}$ of H s.t. $\forall n \in \mathbb{N}$ the set $\{e_1, \dots, e_{k_n}\}$ is an ONB of $P_n(H)$. Then define the sequence

[11] One could have also argued that w.l.o.g. the P_n can be chosen s.t. $P_n \uparrow I$.

$$Z_n := \sum_{i=1}^{k_n} \xi_i \, e_i, \quad n \in \mathbb{N}$$

of H-valued random variables. By the above, for every $n \in \mathbb{N}$

$$\begin{aligned} &\mathbb{P}\left\{\omega \in \Omega : \|Z_{n+1} - Z_n\|_E > 2^{-n}\right\} && (4.9)\\ &= \mathbb{P}\left\{\omega \in \Omega : \left\| \sum_{i=k_n+1}^{k_{n+1}} \xi_i(\omega) e_i \right\|_E > 2^{-n}\right\} \\ &= \nu_H\Big(x \in H : \|(P_{n+1} - P_n)x\|_E > 2^{-n}\Big) < 2^{-n}. && (4.10)\end{aligned}$$

Hence

$$\sum_{n=1}^{\infty} \mathbb{P}\Big\{\omega \in \Omega : \|Z_{n+1} - Z_n\|_E > 2^{-n}\Big\} < \sum_{n=1}^{\infty} 2^{-n} < \infty,$$

which, via the Borel–Cantelli lemma, implies that $\mathbb{P}$-a.s. $\|Z_{n+1} - Z_n\|_E \leq 2^{-n}$ for all but finitely many $n \in \mathbb{N}$. In particular this means that $\mathbb{P}$-a.s. $(Z_n)_{n\in\mathbb{N}}$ is $\|\cdot\|_E$-Cauchy and thus $\mathbb{P}$-a.s. $\|\cdot\|_E$-convergent to a limit $Z := \lim_{n\to\infty} Z_n$.

Define a measure μ on $\mathcal{B}(E)$ as the distribution of Z. Then we want to show that μ is an extension of $\nu_{H*} := \nu_H \circ \mathtt{i}^{-1}$, which is defined on $\mathcal{C}(E)$. Note that since H is embedded in E via the inclusion $\mathtt{i}$, the dual E^* is embedded in $H^* \simeq H$ via the adjoint $\mathtt{i}^*$. We will show that $\widehat{\mu}$ and $\widehat{\nu_{H*}}$ coincide on E^*.

In the following we will denote the action of a linear functional $f \in E^*$ on an element $x \in E$ via $\langle f, x\rangle_E$. Let $f \in E^*$ be arbitrary. Then using the Dominated Convergence Theorem

$$
\begin{aligned}
\widehat{\mu}(f) &= \int_E \exp\Big\{i\big\langle f, x\big\rangle_E\Big\}\ \mathrm{d}\mu(x) \\
&= \int_\Omega \exp\Big\{i\big\langle f, \mathtt{i}Z(\omega)\big\rangle_E\Big\}\ \mathrm{d}\mathbb{P}(\omega) \\
&= \lim_{N\to\infty}\int_\Omega \exp\Big\{i\sum_{n=1}^N \xi_n(\omega)\langle f, \mathtt{i}e_n\rangle_E\Big\}\ \mathrm{d}\mathbb{P}(\omega) \\
&= \lim_{N\to\infty}\prod_{n=1}^N\int_\Omega \exp\Big\{i\xi_n(\omega)\langle f, \mathtt{i}e_n\rangle_E\Big\}\ \mathrm{d}\mathbb{P}(\omega) \\
&= \lim_{N\to\infty}\prod_{n=1}^N\int_\Omega \exp\Big\{i\xi_n(\omega)\langle \mathtt{i}^*f, e_n\rangle_H\Big\}\ \mathrm{d}\mathbb{P}(\omega) \\
&= \lim_{N\to\infty}\prod_{n=1}^N \exp\Big\{-\frac{1}{2}\langle \mathtt{i}^*f, e_n\rangle_H^2\Big\} \\
&= \exp\Big\{-\frac{1}{2}\sum_{n=1}^\infty\langle \mathtt{i}^*f, e_n\rangle_H^2\Big\} \\
&= \exp\Big\{-\frac{1}{2}\|\mathtt{i}^*f\|_H^2\Big\}.
\end{aligned}
$$

Furthermore

$$
\begin{aligned}
\widehat{\nu_{H*}}(f) &= \int_\mathbb{R} \exp\{it\}\ \mathrm{d}\left[\nu_{H*}\circ f^{-1}\right](t) \\
&= \int_\mathbb{R} \exp\{it\}\ \mathrm{d}\left[\nu_H\circ \mathtt{i}^{-1}\circ f^{-1}\right](t) \\
&= \int_\mathbb{R} \exp\{it\}\ \mathrm{d}\left[\nu_H\circ (f\circ\mathtt{i})^{-1}\right](t) \\
&= \int_\mathbb{R} \exp\{it\}\ \mathrm{d}\left[\nu_H\circ (\mathtt{i}^*f)^{-1}\right](t) \\
&= \widehat{\nu_H}\big[\mathtt{i}^*(f)\big] = \exp\left\{-\frac{1}{2}\|\mathtt{i}^*(f)\|_H^2\right\},
\end{aligned}
$$

where in the last line we used Proposition 4.6. This is what was to be shown.
(2) Recall that the Cameron–Martin subspace $\mathbf{H}(\mu)$ consists of all those elements $h\in E$ s.t.

$$
\big(\mathbf{K}(\mu), \langle\cdot,\cdot\rangle_{L^2(E,\mu)}\big)\to\mathbb{R},\quad f\mapsto f(h)
$$

is continuous. In order to show $(H, \|\cdot\|_H) = \big(\mathbf{H}(\mu), \|\cdot\|_{\mathbf{H}(\mu)}\big)$ we show equality as sets and isometry.

Inclusion "$H\subseteq\mathbf{H}(\mu)$": Let $h\in H$ and $f\in E^*$ be arbitrary. Firstly,

$$|\langle f, h\rangle_E| = |\langle \mathtt{i}^* f, h\rangle_H| \le \|\mathtt{i}^* f\|_H \|h\|_H.$$

Secondly,

$$\begin{aligned}
\langle f, f\rangle_{L^2(E,\mu)} &= \int_E |f(x)|^2 \mathrm{d}\mu(x) = \int_\Omega |f\big(\mathtt{i}Z(\omega)\big)|^2 \mathrm{d}\mathbb{P}(\omega) \qquad (4.11)\\
&= \mathrm{Var}\left[f(\mathtt{i}Z)\right] = \mathrm{Var}\left[f(\lim_{n\to\infty} \mathtt{i}Z_n)\right] = \mathrm{Var}\left[\lim_{n\to\infty} f(\mathtt{i}Z_n)\right] = \lim_{n\to\infty} \mathrm{Var}\left[f(\mathtt{i}Z_n)\right]\\
&= \lim_{n\to\infty} \sum_{j=1}^n |\langle f, \mathtt{i}e_j\rangle_E|^2 = \sum_{j=1}^\infty |\langle f, \mathtt{i}e_j\rangle_E|^2\\
&= \sum_{j=1}^\infty |\langle \mathtt{i}^* f, e_j\rangle_H|^2 = \|\mathtt{i}^* f\|_H^2.
\end{aligned}$$

Here we used that $f(Z_n) \to f(Z)$ $\mathbb{P}$-a.s. since f is continuous and thus $\mathrm{Var}\big(f(Z_n)\big) \to \mathrm{Var}\big(f(Z)\big)$ because both $f(Z_n)$ and $f(Z)$ are Gaussian. Hence we have

$$|\langle f, h\rangle_E| \le \langle f, f\rangle_{L^2(E,\mu)} \|h\|_H. \qquad (4.12)$$

Dividing by $\langle f, f\rangle_{L^2(E,\mu)}$ and taking the supremum over all $f \in \mathbf{K}(\mu)$ shows that $h \in \mathbf{H}(\mu)$.

Inclusion "$H \supseteq \mathbf{H}(\mu)$": In order to show the other inclusion we want to show that $(H, \|\cdot\|_H)$ is dense in $\big(\mathbf{H}(\mu), \|\cdot\|_{\mathbf{H}(\mu)}\big)$. Since $(H, \|\cdot\|_H)$ is complete and $\|\cdot\|_{\mathbf{H}(\mu)} = \|\cdot\|_H$ (due to the isometry, which will be shown last) we deduce that $\big(H, \|\cdot\|_{\mathbf{H}(\mu)}\big)$ is complete and hence a closed subspace of $\big(\mathbf{H}(\mu), \|\cdot\|_{\mathbf{H}(\mu)}\big)$. By Theorem 3.16 we know that the adjoint of the inclusion $\mathtt{i}_\mu : \big(\mathbf{H}(\mu), \|\cdot\|_{\mathbf{H}(\mu)}\big) \to E$ is an isometry (1) and has dense image (2). We will show that the orthogonal complement of $H \subseteq \mathbf{H}(\mu)$ is $\{0\}$. Let $g \in \mathbf{H}(\mu)$ s.t. $\langle g, h\rangle_{\mathbf{H}(\mu)} = 0$ for every $h \in H$. Choose a sequence $\mathtt{i}_\mu^*(f_n) \to g$ in the $\|\cdot\|_{\mathbf{H}(\mu)}$-norm (which exists by (2)). Then $(f_n)_{n\in\mathbb{N}} \subseteq E^*$ is $\mathfrak{q}$-Cauchy (by (1)), thus $(\mathtt{i}^* f_n)_{n\in\mathbb{N}} \subseteq E^*$ is $\|\cdot\|_H$-Cauchy (by equation (4.11)), and hence $\|\cdot\|_H$-convergent to some $k \in H$. However, for every $h \in H$

$$\langle k, h\rangle_H \leftarrow \langle \mathtt{i}^*(f_n), h\rangle_H = \langle f_n, \mathtt{i}h\rangle_H = \langle f_n, \mathtt{i}_\mu h\rangle_H = \langle \mathtt{i}_\mu^*(f_n), h\rangle_{\mathbf{H}(\mu)} \to \langle g, h\rangle_{\mathbf{H}(\mu)} = 0$$

when $n \to \infty$. Thus $k = 0$, implying

$$\|g\|_{\mathbf{H}(\mu)}^2 = \lim_{n\to\infty} \|\mathtt{i}_\mu^*(f_n)\|_{\mathbf{H}(\mu)}^2 = \lim_{n\to\infty} \langle f_n, f_n\rangle_{L^2(E,\mu)} = \lim_{n\to\infty} \|\mathtt{i}^*(f_n)\|_H^2 = \|k\|_H^2 = 0,$$

and thus $g = 0$. This shows the claim.

Inequality "$\|\cdot\|_{\mathbf{H}(\mu)} \le \|\cdot\|_H$": Let $h \in H \cap \mathbf{H}(\mu)$ be arbitrary. Rearranging equation (4.12) and taking the supremum over $f \in E^*$ gives the result.

Inequality "$\|\cdot\|_{\mathbf{H}(\mu)} \ge \|\cdot\|_H$": Let $h \in H \cap \mathbf{H}(\mu)$ be arbitrary. Choose a sequence $(f_n)_{n\in\mathbb{N}} \subseteq E^*$ s.t. $\mathtt{i}^* f_n \to h$ in H (which exists since $\mathtt{i}$ is injective, and thus $\mathtt{i}^*$ has dense image). Then by definition we have

$$\frac{|\langle f_n, h\rangle_E|}{\langle f_n, f_n\rangle_{L^2(E,\mu)}} \leq \|h\|_{\mathbf{H}(\mu)}, \quad \forall n \in \mathbb{N}.$$

Now, for the enumerator on the left hand side we have

$$|\langle f_n, h\rangle_E| = \langle \mathtt{i}^* f_n, h\rangle_H \to \|h\|_H^2$$

and for the denominator

$$\langle f_n, f_n\rangle_{L^2(E,\mu)} = \|\mathtt{i}^* f_n\|_H^2 \to \|h\|_H^2,$$

where we used that $\mathtt{i}^*$ is an isometry (see (4.11)). Thus the other inequality is shown. □

Definition 4.15. An **abstract Wiener space** is a quadruple $(E, H, \mathtt{i}, \mu)$ consisting of

(1) a separable Banach space $(E, \|\cdot\|_E)$,

(2) a separable Hilbert space $(H, \langle\cdot,\cdot\rangle_H)$,

(3) a continuous, linear embedding $\mathtt{i} : H \hookrightarrow E$ with dense image s.t. $\|\mathtt{i}(\cdot)\|_E$ is a measurable norm in the sense of Gross,

(4) a Gaussian measure μ on $\mathcal{B}(E)$ which is the extension of $\nu_H \circ \mathtt{i}^{-1}$ from $\mathcal{C}(E)$ to $\mathcal{B}(E)$,

where ν_H is the canonical cylinder measure on H as defined in Definition 4.4.

The next theorem shows that every centred Gaussian measure on a separable Banach space arises in this fashion - it is a converse to Theorem 4.14.

Theorem 4.16. *Let $(E, \|\cdot\|_E)$ be a separable Banach space, μ a Gaussian measure on it, and $\mathbf{H}(\mu)$ its Cameron–Martin space. If $\mathbf{H}(\mu) \hookrightarrow (E, \|\cdot\|_E)$ is a dense embedding, then $(E, \mathbf{H}(\mu), \boldsymbol{i}, \mu)$ is an abstract Wiener space.*

Proof. See [28]. □

Finally, we characterize those AWS for which E is a Hilbert space.

Theorem 4.17 ([8, Prop. 4.59, Thm. 4.60, Cor. 4.62])**.** *Let $(H, \|\cdot\|_H)$ be a separable Hilbert space, ν_H its canonical cylinder measure, and $\|\boldsymbol{i}(\cdot)\|_E$ a norm on H. Then the following are equivalent*

(1) $\|\cdot\|_E$ is a measurable norm w.r.t. (H, ν_H) and induced by an inner product, i.e.

$$\|\boldsymbol{i}h\|_E = \sqrt{\langle \boldsymbol{i}h, \boldsymbol{i}h\rangle_E}, \quad \forall h \in H.$$

(2) $\|\cdot\|_E$ is induced by a positive definite, Hermitian, Hilbert–Schmidt class linear operator A on H, i.e.

$$\|\boldsymbol{i}h\|_E = \|Ah\|_H, \quad \forall h \in H.$$

Proof. "(2) ⇒ (1)" For every $h \in H$ we have

$$\|\mathtt{i}h\|_E = \|Ah\|_H = \sqrt{\langle Ah, Ah\rangle_H}$$

i.e. $\|\cdot\|_E$ is induced by an inner product. To show that $\|\cdot\|_E$ is measurable, let $\varepsilon > 0$. Choose an ONB $(e_n)_{n\in\mathbb{N}}$ of H, choose n_ε s.t.

$$\underbrace{\sum_{n=n_\varepsilon+1}^{\infty} \|Ae_n\|_H^2}_{=\|A|_{(P_\varepsilon H)^\perp}\|_{\mathrm{HS}}^2} < \varepsilon^3, \tag{4.13}$$

and define P_ε as the orthogonal projection onto the subspace spanned by $\{e_1, \dots, e_{n_\varepsilon}\}$. This is possible since A is Hilbert–Schmidt. For any orthogonal projection P s.t. $P \perp P_\varepsilon$ we now have

$$\begin{aligned}\nu_H\Big\{h \in H : \|P_\varepsilon h\|_E > \varepsilon\Big\} &= \nu_H\Big\{h \in H : \|AP_\varepsilon h\|_H > \varepsilon\Big\} \\ = \mathbb{P}\Big\{h \in H : \|A\eta\|_H > \varepsilon\Big\} &\le \frac{\mathbb{E}\Big[\|AE\|_H^2\Big]}{\varepsilon^2},\end{aligned}$$

where η is an H-valued random variable which is standard normally distributed on the subspace spanned by $\{e_1, \dots, e_{n_\varepsilon}\}$ (this is analogous to the argument in formulas (4.9)-(4.10)) and we applied Chebyshev's inequality in the last step. The claim now follows by (4.13) and

$$\begin{aligned}\mathbb{E}\Big[\|A\eta\|_H^2\Big] &= \sum_{i=1}^{n_\varepsilon} \mathbb{E}\Big[|\langle A\eta, e_i\rangle_H|^2\Big] \\ &= \sum_{i=1}^{n_\varepsilon} \mathbb{E}\Big[|\langle \eta, Ae_i\rangle_H|^2\Big] \\ &= \sum_{i=1}^{n_\varepsilon} \|Ae_i\|_H^2 \\ &= \|A|_{PH}\|_{\mathrm{HS}}^2 \le \|A|_{(P_\varepsilon H)^\perp}\|_{\mathrm{HS}}^2\end{aligned}$$

where in the last equality we used that $P \perp P_\varepsilon$.
"(1) ⇒ (2)" Since $\|\cdot\|_E$ is a measurable norm, by Theorem 4.14 there exists an AWS construction. The linear operator $A := (\mathtt{i}^*\mathtt{i})^{\frac{1}{2}}$ satisfies

$$\|\mathtt{i}h\|_E^2 = \langle \mathtt{i}h, \mathtt{i}h\rangle_E = \langle \mathtt{i}^*\mathtt{i}h, h\rangle_E = \langle (\mathtt{i}^*\mathtt{i})^{\frac{1}{2}}h, (\mathtt{i}^*\mathtt{i})^{\frac{1}{2}}h\rangle_E = \langle Ah, Ah\rangle_E,$$

is positive definite since $\mathtt{i}$ and $\mathtt{i}^*$ are injective, is Hermitian, and is Hilbert–Schmidt since $\mathtt{i}^*\mathtt{i}$ is:

$$\begin{aligned}\|\mathtt{i}^*\mathtt{i}\|_{\mathrm{HS}}^2 &= \sum_{n=1}^{\infty}\langle \mathtt{i}^*\mathtt{i}e_n, \mathtt{i}^*\mathtt{i}e_n\rangle_H = \sum_{n=1}^{\infty}\langle \mathtt{i}e_n, \mathtt{i}e_n\rangle_{L^2(E,\mu)} = \sum_{n=1}^{\infty}\int_E |\langle \mathtt{i}e_n, x\rangle_E|^2 \,\mathrm{d}\mu(x) \\ &= \sum_{n=1}^{\infty}\int_E |\langle \mathtt{i}^*x, e_n\rangle_H|^2 \,\mathrm{d}\mu(x) = \int_E \sum_{n=1}^{\infty} |\langle \mathtt{i}^*x, e_n\rangle_H|^2 \,\mathrm{d}\mu(x) = \int_E \|\mathtt{i}^*x\|_H^2 \,\mathrm{d}\mu(x) \\ &\le \|\mathtt{i}^*\|^2 \int_E \|x\|_E^2 \,\mathrm{d}\mu(x) < \infty\end{aligned}$$

where we used Fubini's Theorem to interchange the sum and the integral, and Fernique's Theorem to show that the last expression is finite. □

Remark 4.18. Example 4.10 is an instance of the above theorem.

4.4 Examples

Brownian Motion on $[0,1]$ The most important example of the AWS construction is of course the classical Wiener space. In the cast of Definition 4.15 that is the separable Banach space $\big(C_0[0,1], \|\cdot\|_\infty\big)$, the separable Hilbert space $\big(H_0^1[0,1], \langle\cdot,\cdot\rangle_{H_0^1[0,1]}\big)$, the inclusion $\mathtt{i} : H_0^1[0,1] \hookrightarrow C_0[0,1]$, and the classical Wiener measure which is an extension of the canonical cylinder measure on $\big(H_0^1[0,1], \langle\cdot,\cdot\rangle_{H_0^1[0,1]}\big)$.

Note that what was constructed in subsection 3.3.3 and in section 4.1 really are the same objects, but seen from two different points of view. In subsection 3.3.3 we started with the state space $C_0[0,1]$ and the covariance structure $\mathrm{Cov}(B_s, B_t) = s \wedge t$, $0 \le s,t \le 1$ of Brownian motion, deduced the covariance form $\mathfrak{q}_\mu$, and then computed the Cameron–Martin space $\big(\mathbf{H}(\mu), \langle\cdot,\cdot\rangle_{\mathbf{H}(\mu)}\big)$, which coincided with $\big(H_0^1[0,1], \langle\cdot,\cdot\rangle_{H_0^1[0,1]}\big)$. In section 4.1, we started from an ONB $\big\{s_{n,k} : k \in I(n), n \in \mathbb{N}\big\}$ of the separable Hilbert space $H_0^1[0,1]$ (and thus a canonical cylinder measure $\nu_{H_0^1[0,1]}$) and obtained a measure μ (the distribution of $B = \sum_{n=1}^\infty \sum_{k\in I(n)} \xi_{n,k} s_{n,k}$) on the space $C_0[0,1]$, which is the closure of $H_0^1[0,1]$ w.r.t. the measurable norm $\|\cdot\|_\infty$ (see Proposition 4.13). Theorem 4.14 then revealed that $H_0^1[0,1]$ coincides with the Cameron–Martin space of $(C_0[0,1], \mu)$ and that μ is an extension of $\nu_{H_0^1[0,1]} \circ \mathtt{i}^{-1}$.

Remark 4.19. The fact that the Hilbert spaces of both constructions coincide is part of the theory. The fact that the Banach spaces coincide is merely a result of our choice of state space in subsection 3.3.3 and choice of measurable norm in section 4.1. In subsection 3.3.3 we could have just as well started with the space $C[0,1]$, instead of $C_0[0,1]$, as the state space of Brownian motion, while in section 4.1 we could have chosen a different measurable norm (e.g. via Theorem 4.17) s.t. the resulting Banach space is contained in $C_0[0,1]$.

Brownian Bridge on $[0,1]$ See [8, Ex. 4.34]. A Brownian bridge (tied down at 1) is a centred Gaussian process with covariance structure

$$\mathrm{Cov}(X_s, X_t) = s \wedge t - st, \quad s,t \in [0,1].$$

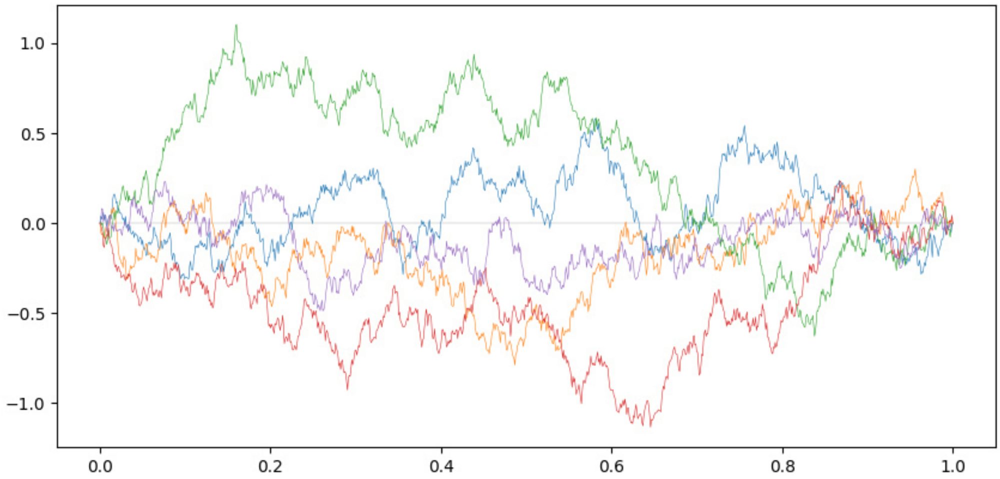

Figure 6: Sample paths of a Brownian bridge (tied down at 1).

It can be written as

$$X_t = B_t - tB_1, \quad t \in [0,1]$$

and therefore has continuous sample paths that start at 0 a.s. The Cameron–Martin space of its distribution is given by those functions $h \in H_0^1[0,1]$ for which the norm induced by

$$\langle h_1, h_2 \rangle := \int_0^1 \left[h_1'(t) + \frac{h_1(t)}{1-t} \right] \left[h_2'(t) + \frac{h_2(t)}{1-t} \right] \mathrm{d}t, \quad h_1, h_2 \in H_0^1[0,1],$$

is finite. Thus, according to Theorem 4.14, the distribution of a Brownian bridge may be constructed as the distribution of

$$\sum_{n=1}^{\infty} \xi_n e_n,$$

where $(e_n)_{n\in\mathbb{N}}$ is an ONB of the Cameron–Martin space and $(\xi_n)_{n\in\mathbb{N}}$ is a sequence of iid $\mathcal{N}(0,1)$-distributed, real-valued random variables.

Fractional Brownian motion on $[0,1]$ A fractional Brownian motion with Hurst parameter $H \in (0,1)$ is a centred Gaussian process with covariance structure

$$\operatorname{Cov}(X_s, X_t) = \frac{1}{2}\big(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\big), \quad s,t \in [0,1].$$

For $H = \frac{1}{2}$ this is standard BM. Its sample paths are continuous (with local Hölder continuity controlled by the parameter H) and start at 0 a.s. The Cameron–Martin space of its distribution and many other aspects are investigated in [5, Chap. 4].

Ornstein–Uhlenbeck process on $[0, 1]$ See [8, Ex. 4.33]. An Ornstein–Uhlenbeck process (starting at 0) is a centred Gaussian process with covariance structure

$$\mathrm{Cov}(X_s, X_t) = \frac{\sigma^2}{2}\Big(\exp\big\{-|s-t|\big\} - \exp\big\{-(t+s)\big\}\Big), \quad s, t \in [0, 1].$$

It satisfies the stochastic ODE

$$\mathrm{d}X_t = -X_t\ \mathrm{d}t + \sigma\ \mathrm{d}B_t, \quad X_0 = 0,$$

where $(B_t)_{t\in[0,1]}$ is a BM and $\sigma > 0$. Its sample paths are continuous and start at 0 a.s. The Cameron–Martin space of its distribution is given by those functions $h \in H_0^1[0, 1]$ for which the norm induced by

$$\langle h_1, h_2\rangle := \frac{1}{\sigma^2}\int_0^1 h_1'(t)h_2'(t) + h_1(t)h_2(t)\mathrm{d}t, \quad h_1, h_2 \in H_0^1[0, 1],$$

is finite.

5 Application to Large Deviations

5.1 Primer on Large Deviations

The following are essentially the introductory remarks of T. Bodineau's talk "Nonequilibrium Statistical Mechanics & Large Deviation Theory" at IHÉS [1].

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of iid, real-valued, and square-integrable random variables with $\mathbb{E}(X_1) = 0$ and $\mathrm{Var}(X_1) = 1$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, by the Strong Law of Large Numbers, the sample averages converge:

$$\frac{1}{N}\sum_{n=1}^{N} X_n \to 0 = \mathbb{E}[X_1], \quad N \to \infty, \quad \mathbb{P}\text{- a.s.},$$

and hence for any $x > 0$

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{n=1}^{N} X_n\right| \geq x\right) \to 0, \quad N \to \infty.$$

In other words, the probability of observing the sample average outside of $(-x, x)$ goes to 0 with an increasing number N of samples. We are looking for a quantitative description of these "large deviations", i.e. we want to study the asymptotic behaviour of $\mathbb{P}\Big(\frac{1}{N}\sum_{n=1}^{N} X_n \geq x\Big)$ for $x \in \mathbb{R}$. An immediate upper bound can be constructed as follows:

Let $(X_n)_{n\in\mathbb{N}}$ be as above with law $\mathcal{L}(X_n) = \mu$ and additionally assume exponential integrability, i.e. $\mathbb{E}\Big[\exp(\lambda X_1)\Big] < \infty$ for every $\lambda \in \mathbb{R}$. Then for any $x \in \mathbb{R}$ and $\lambda > 0$ it is true that

$$\mathbb{P}\left(\frac{1}{N}\sum_{n=1}^{N} X_n \geq x\right) = \mathbb{P}\left(\lambda\sum_{n=1}^{N} X_n \geq \lambda N x\right) = \mathbb{P}\left(\exp\left\{\lambda\sum_{n=1}^{N} X_n\right\} \geq \exp\{\lambda N x\}\right).$$

Via Chebyshev's inequality the last term can be bounded above by

$$\mathbb{E}\left[\exp\left\{\lambda\sum_{n=1}^{N} X_n\right\}\right] e^{-\lambda N x} = \mathbb{E}\Big[e^{\lambda X_n}\Big]^N e^{-\lambda N x}.$$

Writing $\psi_\mu(\lambda) := \mathbb{E}\big[e^{\lambda X_n}\big]$ for the moment generating function of μ we obtain

$$\mathbb{P}\left(\frac{1}{N}\sum_{n=1}^{N} X_n \geq x\right) \leq \exp\Big\{ - N\Big[\lambda x - \ln\big(\psi_\mu(\lambda)\big)\Big]\Big\}.$$

Via the Hölder inequality one can show that $\ln\big(\psi_\mu(\cdot)\big)$ is a convex function. Now, since the above is true for arbitrary $\lambda > 0$, we can improve the bound to

$$\mathbb{P}\left(\frac{1}{N}\sum_{n=1}^{N} X_n \geq x\right) \leq \exp\Big\{ - N\Lambda_\mu^*(x)\Big\},$$

where Λ_μ^* is the Legendre transform of the logarithmic moment generating function $\Lambda_\mu(\cdot) = \ln\big(\psi_\mu(\cdot)\big)$ of μ, given by

$$\Lambda_\mu^*(x) := \sup_{\lambda>0}\Big[\lambda x - \ln\big(\psi_\mu(\lambda)\big)\Big], \quad x \in \mathbb{R}.$$

Since $\Lambda_\mu^* \geq 0$, this implies that the probability of observing sample averages in the event $[x,\infty)$ decays exponentially with $N \in \mathbb{N}$. Rewriting, we have

$$\limsup_{N\to\infty} \frac{1}{N}\ln\left\{\mathbb{P}\left(\frac{1}{N}\sum_{n=1}^{N} X_n \geq x\right)\right\} \leq -\Lambda_\mu^*(x), \quad x > 0.$$

One can show that $-\Lambda_\mu^*(x)$ is also a lower bound, ultimately giving

$$\lim_{N\to\infty} \frac{1}{N}\ln\left\{\mathbb{P}\left(\frac{1}{N}\sum_{n=1}^{N} X_n \geq x\right)\right\} = -\Lambda_\mu^*(x), \quad x > 0.$$

The Legendre transform of the logarithmic moment generating function of μ is convex, lower semi-continuous, $\Lambda_\mu^* \geq 0$, and the normalizing sequence $(N)_{N\in\mathbb{N}}$ goes to infinity. It is a special case of the following general set-up.

Definition 5.1. Let S be a polish space, $(\mu_N)_{N\in\mathbb{N}}$ a sequence of probability measures on S and $(\gamma_N)_{N\in\mathbb{N}} \subseteq \mathbb{R}$ a sequence s.t. $\gamma_N \to \infty$. Then a function $I : S \to [0,\infty]$ is called a **rate function** if

(1) $I(x) \geq 0$ for every $x \in S$,

(2) I is lower semi-continuous, i.e. for every $\alpha \geq 0$ the level set $\{x \in S : I(x) \leq \alpha\}$ is closed.

The rate function I is called **good** if for every $\alpha \geq 0$ the level set $\{x \in S : I(x) \leq \alpha\}$ is compact. The sequence $(\mu_N)_{N\in\mathbb{N}}$ is said to satisfy a **large deviation principle** with **rate function** I and **speed** $(\gamma_N)_{N\in\mathbb{N}}$ if for every $A \in \mathcal{B}(S)$

$$-\inf_{x\in\operatorname{int}(A)} I(x) \leq \liminf_{N\to\infty}\left\{\frac{1}{\gamma_N}\ln\left[\mu_N(A)\right]\right\} \leq \limsup_{N\to\infty}\left\{\frac{1}{\gamma_N}\ln\left[\mu_N(A)\right]\right\} \leq -\inf_{x\in\overline{A}} I(x),$$

where $\operatorname{int}(A)$ and $\overline{A}$ denote the interior and closure of A, respectively.

5.2 Large Deviations on Abstract Wiener Spaces

We would like to study large deviations for families of measures on separable Banach spaces. To give a more concrete sense to "large deviations" there should be an analogue of the Strong Law of Large Numbers, i.e. an event from which atypical observation deviate. This role is played by the Strong Law of Large Numbers of Ranga Rao.

Theorem 5.2 (Ranga Rao). *Let $(E, \|\cdot\|_E)$ be a separable Banach space and $(X_n)_{n\in\mathbb{N}}$ a sequence of iid, μ-distributed, E-valued random variables, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If $\int_E \|x\|_E \,\mathrm{d}\mu(x) < \infty$, then there exists an $\mathfrak{m}_\mu \in E$ s.t.*

$$\frac{1}{N}\sum_{n=1}^{N} X_n \to \mathfrak{m}_\mu \quad \mathbb{P}-a.s.$$

Moreover, $\mathfrak{m}_\mu$ is the unique element in E s.t.

$$f\left(\mathfrak{m}_\mu\right) = \int_E f(x)\,\mathrm{d}\mu(x), \quad f \in E^*,$$

and thus coincides with the mean of μ as in Definition 2.7.

Proof. See [7, Thm. 3.3.4.]. □

The classical theorem of Schilder establishes the existence of a large deviation principle for the distribution of sample averages $(\mu_N)_{N\in\mathbb{N}}$ where μ is the classical Wiener measure on $C_0[0,1]$ (see [6, Thm. 5.2.3.] and see [7, Sec. 1.3] for the case of Brownian motion on $[0,\infty$ which requires a slightly different state space). We want to generalize this to an abstract Wiener measure on a separable Banach space. The proof relies on the Donsker–Varadhan Theorem.

Theorem 5.3. *(Donsker–Varadhan) Let $(E, \|\cdot\|_E)$ be a separable Banach space and $(X_n)_{n\in\mathbb{N}}$ a sequence of iid E-valued random variables with law $\mathcal{L}(X_1) = \mu \in M_1(E)$ s.t.*

$$\int_E \exp\left\{\alpha\|x\|_E\right\}\,\mathrm{d}\mu(x) < \infty, \quad \forall\alpha > 0. \tag{5.1}$$

Then the family $\{\mu_N\}_{N\in\mathbb{N}} \subseteq M_1(E)$, defined by

$$\mu_N = \mathcal{L}\left(\frac{1}{N}\sum_{n=1}^{N} X_n\right), \quad N \in \mathbb{N},$$

satisfies a large deviation principle with good, convex rate function

$$I(x) = \inf_{\substack{\nu \in M_1(E) \\ m(\nu)=x}} \mathbf{H}(\nu \| \mu) = \Lambda_\mu^*(x).$$

where $\mathbf{H}(\cdot\|\mu) : M_1(E) \to \mathbb{R}$ *is the relative entropy functional and* Λ_μ^* *is the Legendre transform of the logarithmic moment generating function.*

Proof. See [7, Thm. 3.3.11. & Ex. 3.3.12.]. □

Theorem 5.4. *(Generalized Schilder) Let* $(E, H, \boldsymbol{i}, \mu)$ *be an abstract Wiener space, where* μ *is centred, and* $(X_n)_{n\in\mathbb{N}}$ *a sequence of iid* E*-valued random variables with law* $\mathcal{L}(X_1) = \mu \in M_1(E)$*. Then the family* $\{\mu_N\}_{N\in\mathbb{N}} \subseteq M_1(E)$*, defined by*

$$\mu_N = \mathcal{L}\left(\frac{1}{N}\sum_{n=1}^{N} X_n\right), \quad N \in \mathbb{N},$$

satisfies a large deviation principle with good, convex rate function

$$I(x) = \begin{cases} \frac{1}{2}\|\boldsymbol{i}^{-1}(x)\|_H^2, & x \in H \\ \infty, & x \notin H \end{cases}. \tag{5.2}$$

Proof. By Fernique's Theorem (3.5), condition (5.1) in the Donsker–Varadhan Theorem is satisfied and thus the existence of I is guaranteed. It is just left to show that I has the claimed form. Firstly, notice that the logarithmic moment generating function of μ is

$$\Lambda_\mu(f) = \ln\left\{\int_E \exp\left\{\langle f, x\rangle_E\right\} \mathrm{d}\mu(x)\right\} = \frac{1}{2}\mathfrak{q}(f, f), \quad f \in E^*,$$

since μ is assumed to be centred. This is analogous to the case of a Gaussian measure on $\mathbb{R}^n$ and the proof is similar to Theorem 3.4. Now let $x \in \mathtt{i}(H) \subseteq E$. Then

$$\begin{aligned}
\Lambda_\mu^*(x) &= \sup_{f\in E^*}\left\{\langle f, x\rangle_E - \Lambda_\mu(f)\right\} \\
&= \sup_{f\in E^*}\left\{\left\langle f, \mathtt{i}\left(\mathtt{i}^{-1}(x)\right)\right\rangle_E - \Lambda_\mu(f)\right\} \\
&= \sup_{f\in E^*}\left\{\left\langle \mathtt{i}^*(f), \mathtt{i}^{-1}(x)\right\rangle_H - \frac{1}{2}\mathfrak{q}(f, f)\right\} && (5.3)\\
&= \sup_{f\in E^*}\left\{\left\langle \mathtt{i}^*(f), \mathtt{i}^{-1}(x)\right\rangle_H - \frac{1}{2}\left\langle \mathtt{i}^*(f), \mathtt{i}^*(f)\right\rangle_H\right\} && (5.4)\\
&= \sup_{g\in H}\left\{\left\langle g, \mathtt{i}^{-1}(x)\right\rangle_H - \frac{1}{2}\langle g, g\rangle_H\right\}. && (5.5)
\end{aligned}$$

From line (5.4) to line (5.5) we used that $\mathtt{i}^*(E)$ lies dense in H since $\mathtt{i}$ is injective, and from line (5.3) to line (5.4) we used the following fact: Let $h \in H$ be arbitrary. Then since μ is centred,

$$\mathfrak{q}\left(f, \overline{h}\right) = f(h) = \langle \mathtt{i}^* f, h\rangle_H = \mathfrak{q}\left(\overline{\mathtt{i}^* f}, \overline{h}\right), \quad f \in E^*.$$

Thus $f = \overline{\mathtt{i}^* f}$ for every $f \in E^*$ and therefore $\mathfrak{C}(f) = \mathtt{i}^*(f)$ for every $f \in E^*$ since $\bar{(\cdot)}$ is an isomorphism between $(\mathbf{K}(\mu), \mathfrak{q})$ and $(H, \langle \cdot, \cdot \rangle_H)$ with inverse $\mathfrak{C}$. Among those $g \in H$ with the same H-norm, the supremum in (5.5) is achieved when $g = \alpha \mathtt{i}^{-1}(x)$ for some $\alpha \in \mathbb{R}$. The α maximizing (5.5) is

$$\operatorname*{argmax}_{\alpha \in \mathbb{R}} \left\{ \left\langle \alpha \mathtt{i}^{-1}(x), \mathtt{i}^{-1}(x) \right\rangle_H - \frac{1}{2} \left\langle \alpha \mathtt{i}^{-1}(x), \alpha \mathtt{i}^{-1}(x) \right\rangle_H \right\} = \operatorname*{argmax}_{\alpha \in \mathbb{R}} \left\{ \alpha - \frac{1}{2}\alpha^2 \right\} = 1.$$

Hence $\Lambda_\mu^*(x) = \frac{1}{2}\|\mathtt{i}^{-1}(x)\|_H^2$ for every $x \in \mathtt{i}(H) \subseteq E$. Now assume $x \notin \mathtt{i}(H)$. Then, due to the fact that H coincides with the Cameron–Martin space of (E, μ) (see Theorem 4.14 (2)), we have

$$\begin{aligned} \Lambda_\mu^*(x) &= \sup \left\{ \langle f, x \rangle_E - \Lambda_\mu(f) : f \in E^* \right\} \\ &\geq \sup \left\{ \langle f, x \rangle_E - \frac{1}{2}\mathfrak{q}(f, f) : f \in E^*, \mathfrak{q}(f, f) = 1 \right\} \\ &= \infty. \end{aligned}$$

□

A Remarks on Degeneracy and Support

In the following we will only consider the case of centred Gaussian measures on separable Banach spaces. As mentioned in Remark 3.13, a (highly) degenerate Gaussian measure can make the theory trivial, which is why it is often assumed that Gaussian measures are non-degenerate. However, there are interesting examples of (mildly) degenerate Gaussian measures such as the distribution of a Brownian bridge on $[0, 1]$. The following remarks, largely the ones given in [8, Sec. 4], will show that the degenerate case can essentially be reduced to the non-degenerate case.

Proposition A.1. *Let $(E, \|\cdot\|_E)$ be a separable Banach space and μ a degenerate centred Gaussian measure on $\sigma(E; E^*)$. Then there exists a closed subspace $E_0 \subseteq E$ s.t.*

(1) $(E_0, \|\cdot\|_E)$ is a separable Banach space,

(2) $\mu(E_0) = 1$,

(3) $\mu_0 := \mu|_{E_0}$ is a non-degenerate centred Gaussian measure on $\mathcal{B}(E_0)$,

(4) the Cameron–Martin space $\mathbf{H}(\mu)$ of (E, μ) is contained in E_0 and coincides with the Cameron–Martin space $\mathbf{H}(\mu_0)$ of (E_0, μ_0).

Proof. Define

$$E_0 := \bigcap_{\substack{f \in E^* \\ f \in \ker \mathfrak{q}}} \ker f. \tag{A.1}$$

(1) Since every $f \in E^*$ is a bounded linear functional the kernel $\ker f$ is closed in E, hence the intersection is closed and thus E_0 is closed in E. Thus E_0 is also a separable Banach

space.

(2) Note that $f \in \ker \mathfrak{q}$ implies that $f = 0$ μ-a.s. since

$$\mathfrak{q}(f,f) = \int_E |f|^2 \mathrm{d}\mu = 0$$

So each set in the intersection on the right hand side of equation (A.1) has measure 1. Hence it is sufficient to show that the intersection is countable. Since E is separable, the weak* topology on the unit ball of its dual is also separable and metrizable.[12] In particular, $\ker \mathfrak{q} \cap B_1^{\text{weak}^*}(0)$ has a countable dense subset $\{f_n\}_{n\in\mathbb{N}}$. We now claim

$$E_0 = \bigcap_{n\in\mathbb{N}} \ker f_n. \tag{A.2}$$

The inclusion "$\subseteq$" is clear. For the other direction, let $x \in E$ be s.t. $f_n(x) = 0$ for every $n \in \mathbb{N}$ and let $f \in \ker \mathfrak{q}$ be arbitrary. Since $\{f_n\}_{n\in\mathbb{N}}$ lies dense in $\ker \mathfrak{q} \cap B_1^{\text{weak}^*}(0)$, we may choose a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ s.t. f_{n_k} converges to $\frac{f}{\|f\|_{E^*}}$ in the weak* topology when $k \to \infty$. Then in particular $0 = f_{n_k}(x) \to f(x)$ and hence $f(x) = 0$. Thus x also lies in the left hand side of equation (A.2), which is what was to be shown.

(3) Note firstly that the covariance form $\mathfrak{q}_0$ on E_0^* induced by $\mu|_{E_0}$ agrees with the restriction of $\mathfrak{q}$ to E_0. That is, for every $f_0 \in E_0^*$ and $f \in E^*$ s.t. $f|_{E_0} = f_0$ we have

$$\mathfrak{q}_0(f_0,f_0) = \int_{E_0} |f_0|^2 \mathrm{d}\mu_0 = \int_E |f|^2 \mathrm{d}\mu = \mathfrak{q}(f,f),$$

since $\mu(E_0) = 1$ and thus $\mu(E \setminus E_0) = 0$. Now assume $f_0 \in E_0^*$ s.t. $\mathfrak{q}_0(f_0,f_0) = 0$, use the Hahn–Banach Theorem to choose a bounded extension $f \in E^*$, and note that $\mathfrak{q}(f,f) = 0$. By definition $E_0 \subseteq \ker f$, so f vanishes on all of E_0, and hence $f_0 = f|_{E_0}$ is the zero functional, which proves the claim. To see that $\mu|_{E_0}$ is Gaussian, for any $f_0 \in E_0^*$ use the Hahn–Banach Theorem to choose an extension of f_0 to all of E, use that $\mu(E \setminus E_0) = 0$, and that μ is Gaussian.

(4) Let $h \in E$ s.t. $\|h\|_{\mathbf{H}(\mu)} < \infty$. Then, since μ is assumed to be centred, for any $f \in E^*$ with $\mathfrak{q}(f,f) = 0$ we have $f(h) = 0$ since otherwise

$$\|h\|_{\mathbf{H}(\mu)} = \sup\left\{\frac{|f(h)|}{\sqrt{\mathfrak{q}(f,f)}} : f \in E^*, f \neq 0\right\} \geq \frac{|f(h)|}{\sqrt{\mathfrak{q}(f,f)}} = \infty.$$

Thus $h \in E_0$ by definition of E_0. It is left to show that for any $h \in H$ we have $\|h\|_{\mathbf{H}(\mu)} < \infty$ if and only if $\|h\|_{\mathbf{H}(\mu_0)} < \infty$. For the direction "$\Rightarrow$" restrict functionals in E to functionals on E_0 (in which $\mathbf{H}(\mu)$ lies). For the other direction use Hahn–Banach to extend any functional on E_0 to one on E. $\square$

[12] [22, Thm. 2.6.20] shows that the unit ball in the dual of a normed space is metrizable. By the Banach–Alaoglu Theorem the unit ball is weak* compact. Thus, as a compact metric space, [14, Thm. 5.8] shows that unit ball is separable.

Moreover, E_0, as defined above, coincides with the topological support of μ and, if μ is non-degenerate, then $\ker \mathfrak{q} = 0$ and thus $E_0 = E$.

Proposition A.2. *Let $(E, \|\cdot\|_E)$ be a separable Banach space and μ a centred Gaussian measure on $\sigma(E; E^*)$. Then*

$$\operatorname{supp} \mu = E_0 := \bigcap_{f \in \ker \mathfrak{q}} \ker f, \tag{A.3}$$

where $\operatorname{supp} \mu$ *denotes the weak support of μ, i.e. the smallest weakly closed subset with full measure.*

Proof. Since for all $f \in \ker \mathfrak{q}$ we have $f = 0$ μ-a.s., each set in the intersection on the right hand side of (A.3) has full measure and is weakly closed because f is weakly continuous. Hence "$\subseteq$" is clear. For the other direction, recall that the topological support of a measure can be characterized as the set of points $x \in E$ s.t. each open neighbourhood has positive measure. We will show that for any x in the right hand side of (A.3) and any $\varepsilon > 0$ and $f_1, \dots, f_n \in E^*$ the set

$$B_{\varepsilon, f_1, \dots, f_n}(x) := \bigcap_{i=1}^{n} f_i^{-1}\Big(B_\varepsilon\big(f_i(x)\big)\Big) = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}^{-1} \Big(\bigtimes_{i=1}^{n} \big(f_i(x) - \varepsilon, f_i(x) + \varepsilon\big)\Big) \tag{A.4}$$

has positive measure. Since the sets of the form (A.4) form a basis of the weak topology and thus any open neighbourhood of x needs to contain a set of that form, the monotonicity of μ gives the result. Assume w.l.o.g. that $f_1, \dots, f_k \in \ker \mathfrak{q}$ for some $0 \le k \le n$. Then by the assumption of x being contained in the right hand side of (A.3) and $(f_1, \dots, f_n)$ being jointly Gaussian with distribution $\mathcal{N}(0, \Sigma)$ and $f_1, \dots, f_k$ degenerate, we have

$$\begin{aligned}
\mu(B_{\varepsilon, f_1, \dots, f_n}) &= \mu\left(\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}^{-1} \left(\bigtimes_{i=1}^{n} (f_i(x) - \varepsilon, f_i(x) + \varepsilon)\right)\right) \\
&= \int_{\left(0, \dots, 0, f_{k+1}(x), \dots, f_n(x)\right) + (-\varepsilon, \varepsilon)^n} \mathrm{d}\mathcal{N}(0, \Sigma) \\
&= \int_{f_n(x)-\varepsilon}^{f_n(x)+\varepsilon} \dots \int_{f_{k+1}(x)-\varepsilon}^{f_{k+1}(x)+\varepsilon} f_{\mathcal{N}(0,\Sigma')} \, \mathrm{d}x_{k+1} \, \dots \, \mathrm{d}x_n > 0,
\end{aligned}$$

where Σ' is the $(n-k) \times (n-k)$-block matrix in Σ corresponding to the non-degenerate entries and $f_{\mathcal{N}(0,\Sigma')}$ is the density of $\mathcal{N}(0, \Sigma')$. $\square$

From the above two propositions, we can conclude that the Cameron–Martin space is contained in the support of μ. Moreover, the support of μ is the $\|\cdot\|_E$-norm-closure of its associated Cameron–Martin space since H lies dense in E, E_0 is a closed subspace of E, and $H \subseteq E_0$.

B Notation

Acronyms & Abbreviations	
TVS	topological vector space
BM	Brownian motion
iid	independent and identically distributed
AWS	abstract Wiener space
ONB	orthonormal basis
s.t.	such that
w.r.t.	with respect to
w.l.o.g.	without loss of generality
loc. co.	locally convex
i.e.	id est (in other words)
e.g.	exempli gratia (for example)
z.B.	zum Beispiel
a.e./a.s.	almost every/ almost surely
sep.	separable

$\mathbb{R}^n$, $n \geq 0$	n-dimensional real space
E'	algebraic dual space of a vector space E, i.e. the vector space of linear functionals $E \to \mathbb{R}$
$(E,\tau)^*$	continuous dual space of a TVS (E,τ), i.e. the vector space of continuous linear functionals $(E,\tau) \to \mathbb{R}$
(E,τ)	locally convex TVS over $\mathbb{R}$; depending on the context a Banach space or a Hilbert space
$\mathbf{M}(\Omega,\mathcal{F})$	space of finite signed measures on a measurable space $(\Omega,\mathcal{F})$
$C(\Omega,\tau)$, $\big(C_x(\Omega,\tau)\big)$	space of continuous real-valued functions on a topological space (Ω,τ) (starting at $x \in \Omega$)
$H^1[0,1]$, $\big(H_0^1[0,1]\big)$	first Hilbert-Sobolev space on the unit interval $[0,1]$ (of functions starting at 0)
$\sigma(\mathcal{A})$	smallest sigma-algebra containing the family of sets $\mathcal{A}$, i.e. the sigma-algebra generated by $\mathcal{A}$
$\sigma(E;F)$	smallest sigma-algebra on E making all functionals $E \to \mathbb{R}$ in F measurable
$\mathcal{B}(E)$	Borel sigma-algebra on the topological space (E,τ)
$\mathcal{C}(E)$	algebra of cylinder sets of a locally convex TVS (E,τ)
Λ_μ / Λ_μ^*	logarithmic generating function of a measure μ and its Legendre transform
ev_t	evaluation functional at t, which acts on a function via $f \mapsto f(t)$
ϕ_X	characteristic function of a real-valued random variable X
ψ_X / ψ_μ	moment generating function of a real-valued random variable X or a measure μ
$X \sim \mu$, $\mu = \mathcal{L}(X)$	the random variable X has distribution μ
$\\|\cdot\\|_{\mathrm{TV}}$	total variation norm
$\mathcal{N}(\mu,\Sigma)$	Gaussian measure with expectation μ and covariance matrix Σ
ν_H	canonical cylinder measure of a separable Hilbert space $(H,\langle\cdot,\cdot\rangle_H)$

References

[1] Thierry Bodineau. Nonequilibrium statistical mechanics & large deviation theory, 2016. Institut des Hautes Études Scientifiques (IHÉS), https://www.youtube.com/watch?v=s4pznYlf0aM.

[2] Vladimir I. Bogachev. *Gaussian Measures*, volume 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, 1998.

[3] Robert Brown. A brief Account of Microscopical Observations made in the Months of June, July and August, 1827, on the Particles contained in the Pollen of Plants; and on the general Existence of active Molecules in Organic and Inorganic Bodies. *The Philosophical Magazine of Annals of Philosophy*, pages 161–173, 1828.

[4] John B. Conway. *A Course in Functional Analysis*. Springer, 2nd edition, 1990.

[5] Laurent Decreusefond. *Selected Topics in Malliavin Calculus*, 2019. Université Paris-Saclay, Lecture notes.

[6] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*. Springer, 2009.

[7] Jean-Dominique Deuschel and Daniel W. Stroock. *Large Deviations*. Academic Press Inc., 1989.

[8] Nathaniel Eldredge. *Analysis and Probability on Infinite-Dimensional Spaces*, 2016. arXiv: 1607.03591v2, Lecture Notes.

[9] Jürgen Elstrodt. *Maß- und Integrationstheorie*. Springer Spektrum, 8th edition, 2018.

[10] Xavier Fernique. *Intégrabilité des vecteurs gaussiens*, pages 1698–1699. CR Acad. Sci. Paris Serie A, 1970.

[11] Richard Feynman. *The Principle of Least Action in Quantum Mechanics*. PhD thesis, Princeton University, 1942.

[12] Gerald B. Folland. *A Course in Abstract Harmonic Analysis*. Taylor & Francis Group, 2015.

[13] Peter K. Friz. *An Introduction to Malliavin Calculus*, 2002. Lecture Notes.

[14] Theodore W. Gamelin and Robert Everist Greene. *Introduction to Topology*. Dover, 1999.

[15] Leonard Gross. Abstract Wiener Spaces. In *Proc. of the 5th Berkley Symposium on Mathematical Statistics and Probability*. Springer, 1967.

[16] Martin Hairer. *An Introduction to Stochastic PDEs*, 2009. arXiv:0907.4178, Lecture Notes.

[17] Brian C. Hall. *Quantum Theory for Mathmaticians*. Springer, 2013.

[18] Zhi-yuan Huang and Jian-an Yan. *Introduction to Infinite Dimensional Stochastic Analysis*. Springer Science & Business Media, 2000.

[19] Gopinath Kallianpur. *Stochastic Filtering Theory*, volume 13. Springer Science & Business Media, 2013.

[20] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, 1998.

[21] Alexander S. Kechris. *Classical Descriptive Set Theory.* Springer-Verlag, 1995.

[22] Robert E. Megginson. *An Introduction to Banach Space Theory.* Springer Science & Business Media, 1998.

[23] Uttara Naik-Nimbalkar. *Bochner Property in Banach Spaces.* Gauthier-Villars, 1981.

[24] Bernt Øksendal. *Stochastic Differential Equations.* Springer, 2010.

[25] Thanu Padmanabhan. *Quantum Field Theory.* Springer, 2016.

[26] Claudia Prevot and Michael Röckner. *A Concise Course on Stochastic Partial Differential Equations*, volume 13. Springer Science & Business Media, 2013.

[27] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness.* Harcourt Brace Jovanovich, 1975.

[28] Hiroshi Sato. Gaussian Measure on a Banach Space and Abstract Wiener Measure. *Nagoya Math. Journal*, 36, 1969.

[29] Helmut H. Schaefer. *Topological Vector Spaces.* Springer, 5th edition, 1986.

[30] N.N. Vakhania, V.I. Tarieladze, and S.A. Chobanyan. *Probability Distributions on Banach Spaces.* D. Reidel Publishing Company, 1987.

[31] Dirk Werner. *Funktionalanalysis.* Springer, 2011.

[32] Norbert Wiener. Differential space. *J. Math. & Phys.*, page 131–174, 1923.