

MA482 STOCHASTIC ANALYSIS

Lectured by Prof. David Elworthy (kde@maths.warwick.ac.uk)

Typed by Tim Sullivan*(T.J.Sullivan@warwick.ac.uk)

University of Warwick

Term 1, 2005–2006

Printed April 28, 2006

*Thanks to P.H.D. O'Callaghan for pointing out and correcting typographical errors.

Contents

1	Re-Cap of Measure Theory	4
2	Fourier Transforms of Measures	6
3	Gaussian Measures on Finite-Dimensional Spaces	9
3.1	Gaussian Measures	9
3.2	Fourier Transforms of Gaussian Measures	11
4	Gaussian Measures on Banach Spaces	13
5	Cylinder Set Measures	14
6	The Paley-Wiener Map and the Structure of Gaussian Measures	20
6.1	Construction of the Paley-Wiener Integral	20
6.2	The Structure of Gaussian Measures	23
7	The Cameron-Martin Formula: Quasi-Invariance of Gaussian Measures	27
8	Stochastic Processes and Brownian Motion in \mathbb{R}^n	32
8.1	Stochastic Processes	32
8.2	Construction of Itô's Integral	33
9	Itô Integrals as Divergences	48
9.1	The Clark-Ocone Theorem and Integral Representation	48
9.2	Chaos Expansions	56

References

[BP] Czesław Bessaga & Aleksander Pełczyński. *Selected Topics in Infinite-Dimensional Topology*. Warszawa: PWN, 1975.

[MX] Mao Xuerong. *Stochastic Differential Equations and their Applications*. Chichester: Horwood, 1997.

[Ø] Bernt Øksendal, *Stochastic Differential Equations* Sixth Edition. Berlin; New York: Springer, 2003.

[RS] Michael Reed & Barry Simon. *Methods of Modern Mathematical Physics*. New York; London: Academic Press, 1980.

[RW1] L.C.G. Rogers & David Williams. *Diffusions, Markov Processes and Martingales. Vol. 1: Foundations*. Chichester: Wiley, 1994.

[RW2] L.C.G. Rogers & David Williams. *Diffusions, Markov Processes and Martingales. Vol. 2: Itô Calculus*. Chichester: Wiley, 1994.

Introduction

Stochastic analysis can be viewed as a combination of infinite-dimensional analysis, measure theory and linear analysis. We consider spaces such as the infinite-dimensional space of paths in a given state space.

- Around 1900, Norbert Wiener (1894–1964) introduced the notion of Wiener measure. This leads to ideas of “homogeneous chaos” and analysis of brain waves.
- Richard Feynman (1918–1988) worked on quantum mechanics and quantum physics, using ideas like $\int_{\text{paths in } \mathbb{R}^3}$ and $\int_{\text{maps } \mathbb{R}^p \rightarrow \mathbb{R}^q}$.
- Stephen Hawking (1942–) took this idea further: $\int_{\text{universes}}$.
- Edward Witten (1951–) applied these methods to topology, topological invariants and knot theory.
- This area’s connections with probability theory lend it the label “stochastic”. Areas of interest include Brownian motion and other stochastic dynamical systems. A large area of application is mathematical finance.

This course is *not* on stochastic dynamical systems.

1 Re-Cap of Measure Theory

Definitions 1.1. A *measurable space* is pair $\{\Omega, \mathcal{F}\}$ where Ω is a set and \mathcal{F} is a σ -algebra on Ω , so $\emptyset \in \mathcal{F}$, $A \in \mathcal{F} \implies \Omega \setminus A \in \mathcal{F}$ and $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Example 1.2. Given a topological space X , the Borel σ -algebra $\mathcal{B}(X)$ is defined to be the smallest σ -algebra containing all open sets in X . We will always use this unless otherwise stated.

Definitions 1.3. If $\{X, \mathcal{A}\}$ and $\{Y, \mathcal{B}\}$ are measurable spaces, $f : X \rightarrow Y$ is *measurable* if $B \in \mathcal{B} \implies f^{-1}(B) \in \mathcal{A}$. In general, $\sigma(f) := \{f^{-1}(B) | B \in \mathcal{B}\}$ is a σ -algebra on X , the σ -algebra generated by f . It is the smallest σ -algebra on X such that f is measurable into $\{Y, \mathcal{B}\}$.

If X, Y are topological spaces then $f : X \rightarrow Y$ continuous implies that f is measurable — this is not quite trivial.

Definitions 1.4. A *measure space* is a triple $\{\Omega, \mathcal{F}, \mu\}$ with $\{\Omega, \mathcal{F}\}$ a measurable space and μ a *measure* on it, i.e. a $\mu : \mathcal{F} \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$ such that $\mu(\emptyset) = 0$ and $A_1, A_2, \dots \in \mathcal{F}$ disjoint $\implies \mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i) \leq \infty$. The space has *finite measure* (or μ is a *finite measure*) if $\mu(\Omega) < \infty$. $\{\Omega, \mathcal{F}, \mu\}$ is a *probability space* (or μ is a *probability measure*) if $\mu(\Omega) = 1$.

Definition 1.5. Given a measure space $\{\Omega, \mathcal{F}, \mu\}$, a measurable space $\{X, \mathcal{A}\}$ and $f : \Omega \rightarrow X$ measurable, define the *push-forward measure* $f_*\mu$ on $\{X, \mathcal{A}\}$ by $(f_*\mu)(A) := \mu(f^{-1}(A))$. As an exercise, check that this is indeed a measure on $\{X, \mathcal{A}\}$.

Exercises 1.6. (i) Check that $f_*\mu$ is indeed a measure on $\{X, \mathcal{A}\}$.

(ii) Show that if $f, g : \Omega \rightarrow X$ are measurable and $f = g$ μ -almost everywhere, then $f_*\mu = g_*\mu$.

Examples 1.7. (i) Lebesgue measure λ^n on \mathbb{R}^n : Borel measure such that $\lambda^n(\text{rectangle}) = \text{product of side lengths}$, e.g. $\lambda^1([a, b]) = b - a$. This determines λ^n uniquely — see *MA359 Measure Theory*.

Take $v \in \mathbb{R}^n$. Define $T_v : \mathbb{R}^n \rightarrow \mathbb{R}^n$ by $T_v(x) := x + v$. Then $(T_v)_*(\lambda^n) = \lambda^n$ since translations send rectangles to congruent rectangles and λ^n is unique. Thus λ^n is *translation-invariant*.

(ii) Counting measure c on any $\{X, \mathcal{A}\}$: $c(A) := \#A$. Counting measure on \mathbb{R}^n is also translation-invariant.

(iii) Dirac measure on any $\{X, \mathcal{A}\}$: given $x \in X$, define

$$\delta_x(A) := \begin{cases} 1 & x \in A, \\ 0 & x \notin A. \end{cases}$$

Dirac measure δ_x on \mathbb{R}^n is not translation-invariant.

Definitions 1.8. Let X be a topological space (with its usual σ -algebra $\mathcal{B}(X)$). A measure μ on X is *locally finite* if for all $x \in X$, there exists an open $U \subseteq X$ with $x \in U$ and $\mu(U) < \infty$. μ on X is called *strictly positive* if for all non-empty open $U \subseteq X$, $\mu(U) > 0$.

Examples 1.9. (i) λ^n is locally finite and strictly positive.

(ii) c is not locally finite in general, but is strictly positive on any X .

(iii) δ_x is finite, and so locally finite, but is not strictly positive in general.

Proposition 1.10. *Suppose that H is a (separable) Hilbert space with $\dim H = \infty$. Then there is no locally finite translation invariant measure on H except $\mu \equiv 0$. (Therefore, there is no “Lebesgue measure” for infinite-dimensional Hilbert spaces.)*

Recall that

- a topological space X is *separable* if it has a countable dense subset, i.e. $\exists x_1, x_2, \dots \in X$ such that $X = \overline{\{x_1, x_2, \dots\}}$;
- if a metric space X is separable then for any open cover $\{U_\alpha\}_{\alpha \in A}$ of X there exists a countable subcover;
- non-separable spaces include $\mathbb{L}(E; F) := \{\text{continuous linear maps } E \rightarrow F\}$ when E, F are infinite-dimensional Banach spaces, and the space of Hölder-continuous functions $C^{0+\alpha}([0, 1]; \mathbb{R})$.

Proof. Suppose that μ is locally finite and translation invariant. Local finiteness implies that there is an open non-empty U such that $\mu(U) < \infty$. Since U is open, there exist $x \in U$ and $r_0 > 0$ such that $B_{r_0}(x) \subseteq U$. Then $\mu(B_{r_0}(x)) < \infty$ as well. By translation, $\mu(B_r(y)) < \infty$ for all $y \in H$ and $r \leq r_0$. Fix an $r \in (0, r_0)$; then $H = \bigcup_{y \in H} B_r(y)$, an open cover. By separability, there exist $y_1, y_2, \dots \in H$ such that $H = \bigcup_{j=1}^{\infty} B_r(y_j)$, so $\mu(B_r(y_j)) > 0$ for some j , and so $\mu(B_r(y)) > 0$ for all $y \in H$ and $r > 0$. Set $c := \mu(B_{r_0/30}(y))$ for any (i.e., all) $y \in H$. Observe that if e_1, e_2, \dots is an orthonormal basis for H then $B_{r_0/30}(e_j/2) \subseteq B_{r_0}(0)$ for all j . By Pythagoras, these balls are disjoint. $\mu(B_{r_0}(0)) \geq \sum_{j=1}^{\infty} c = \infty$ unless $\mu \equiv 0$. But $\mu \not\equiv 0 \implies \mu(B_{r_0}(0)) < \infty$ by local finiteness, a contradiction. \square

Definition 1.11. Two measures μ_1, μ_2 on $\{\Omega, \mathcal{F}\}$ are *equivalent* if $\mu_1(A) = 0 \iff \mu_2(A) = 0$. If so, write $\mu_1 \approx \mu_2$.

Example 1.12. Standard Gaussian measure γ^n on \mathbb{R}^n :

$$\gamma^n(A) := (2\pi)^{-n/2} \int_A e^{-\|x\|^2/2} dx$$

for $A \in \mathcal{B}(\mathbb{R}^n)$. Here $dx = d\lambda^n(x)$ and $\|x\|^2 = x_1^2 + \dots + x_n^2$ for $x = (x_1, \dots, x_n)$. $\lambda^n \approx \gamma^n$ since $e^{-\|x\|^2/2} > 0$ for all $x \in \mathbb{R}^n$

Definition 1.13. Given a measure space $\{\Omega, \mathcal{F}, \mu\}$, let $f : \Omega \rightarrow \Omega$ be measurable. Then μ is *quasi-invariant under f* if $f_*\mu \approx \mu$, i.e. $\mu(f^{-1}(A)) = 0 \iff \mu(A) = 0$ for all $A \in \mathcal{F}$.

Example 1.14. γ^n is quasi-invariant under all translations of \mathbb{R}^n .

Theorem 1.15. *If E is a separable Banach space and μ is a locally finite Borel measure on E that is quasi-invariant under all translations then either $\dim E < \infty$ or $\mu \equiv 0$.*

The proof of this result is beyond the scope of this course, although it raises the question: are there any “interesting” and “useful” measures on infinite-dimensional spaces?

2 Fourier Transforms of Measures

Definition 2.1. Let μ be a probability measure on a separable Banach space E . Let $E^* := \mathbb{L}(E; \mathbb{R})$ be the dual space. The *Fourier transform* $\hat{\mu} : E^* \rightarrow \mathbb{C}$ is given by

$$\hat{\mu}(\ell) := \int_E e^{i\ell(x)} d\mu(x)$$

for $\ell \in E^*$, where $i = \sqrt{-1} \in \mathbb{C}$. It exists since $\int_E |e^{i\ell(x)}| d\mu(x) = \int_E d\mu(x) = \mu(E) = 1 < \infty$. In fact, for all $\ell \in E^*$, $|\hat{\mu}(\ell)| \leq 1$ and $\hat{\mu}(0) = \mu(E) = 1$.

For E a Hilbert space H with inner product $\langle \cdot, \cdot \rangle$, the Riesz Representation Theorem gives an isomorphism $H^* \rightarrow H : \ell \mapsto \ell^\sharp$, where $\ell^\sharp \in H$ has $\langle \ell^\sharp, x \rangle = \ell(x)$ for all $x \in H$. Therefore, we can consider $\hat{\mu}^\sharp : H \rightarrow \mathbb{C}$ given by

$$\hat{\mu}^\sharp(h) := \int_H e^{i\langle h, x \rangle} d\mu(x).$$

So $\hat{\mu}^\sharp(\ell^\sharp) = \hat{\mu}(\ell)$. Without confusion we write $\hat{\mu}$ for $\hat{\mu}^\sharp$, and so use $\hat{\mu} : H \rightarrow \mathbb{C}$ or $\hat{\mu} : H^* \rightarrow \mathbb{C}$ as convenient.

Example 2.2. For \mathbb{R}^n , if $\mu = f_*\lambda^n$, $f : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$, so $\mu(A) = \int_A f(x) dx$, $f \in L^1$, $\int_{\mathbb{R}^n} f(x) dx = 1$, so $\mu(\mathbb{R}^n) = 1$. Then

$$\hat{\mu}(h) = \int_{\mathbb{R}^n} e^{i\langle h, x \rangle_{\mathbb{R}^n}} d\mu(x) = \int_{\mathbb{R}^n} e^{i\langle h, x \rangle_{\mathbb{R}^n}} f(x) dx,$$

the Fourier transform of f up to signs and constants.

Example 2.3. For a general separable Banach space E , $\mu = \delta_{x_0}$ for some $x_0 \in E$, $\hat{\mu} : E^* \rightarrow \mathbb{C}$ is $\hat{\mu}(\ell) = \int_E e^{i\ell(x)} d\delta_{x_0}(x) = e^{i\ell(x_0)}$. In Hilbert space notation, if $E = H$, we get $\hat{\mu}(h) = e^{i\langle h, x_0 \rangle_H}$.

Proposition 2.4. (Transformation of Integrals.) *Given $\{X, \mathcal{A}, \mu\}$, $\{Y, \mathcal{B}\}$ and $\theta : X \rightarrow Y$ measurable, giving $\theta_*\mu$ on Y , let $f : Y \rightarrow \mathbb{R}$ be measurable. Then $\int_X f \circ \theta d\mu = \int_Y f d(\theta_*\mu)$, in the sense that if one exists then so does the other and there is equality.*

$$\begin{array}{ccc} X & \xrightarrow{\theta} & Y \\ & \searrow f \circ \theta & \downarrow f \\ & & \mathbb{R} \end{array}$$

Proof. By the definition of $\theta_*\mu$ this is true for characteristic functions $f = \chi_B$, $B \in \mathcal{B}$.

$$\begin{aligned} \int_X \chi_B \circ \theta d\mu &= \int_X \chi_{\{x | \theta(x) \in B\}} d\mu \\ &= \mu(\theta^{-1}(B)) \\ &= \theta_*\mu(B) \\ &= \int_Y \chi_B d(\theta_*\mu) \end{aligned}$$

Therefore the claim holds for simple f , and so for measurable f by the approximation definition of the integral. \square

Remark 2.5. Back to $\hat{\mu}$: for a probability measure on a separable Banach space E and $\ell \in E^*$ we have a measure $\mu_\ell := \ell_*\mu$ on \mathbb{R} , and $x \mapsto e^{i\ell(x)}$ factorizes as

$$\begin{array}{ccc} E & \xrightarrow{\ell} & \mathbb{R} \\ & \searrow e^{i\ell(\cdot)} & \downarrow t \mapsto e^{it} \\ & & \mathbb{C} \end{array}$$

$$\begin{aligned} \hat{\mu}(\ell) &:= \int_E e^{i\ell(x)} d\mu(x) \\ &= \int_{\mathbb{R}} e^{it} d\mu_\ell(t) \\ &= \widehat{\mu}_\ell(1) \end{aligned}$$

since $\langle s, t \rangle_{\mathbb{R}} = st$, $\langle 1, t \rangle_{\mathbb{R}} = t$ in the integrand above. Thus $\hat{\mu}$ is determined by $\{\mu_\ell | \ell \in E^*\}$ by the formula $\hat{\mu}(\ell) = \widehat{\mu}_\ell(1)$.

Remark 2.6. Let $T \in \mathbb{L}(E; F)$, E, F separable Banach spaces, μ a probability measure on E , then if $\ell \in F^*$,

$$\begin{aligned} \widehat{T_*\mu}(\ell) &:= \int_F e^{i\ell(y)} d(T_*\mu)(y) \\ &= \int_E e^{i(\ell \circ T)(x)} d\mu(x) \\ &= \hat{\mu}(T^*(\ell)), \end{aligned}$$

where $T^* \in \mathbb{L}(F^*; E^*)$ is the adjoint of T given by $T^*: \ell \mapsto \ell \circ T$.

We ask:

- Can any function $f : E \rightarrow \mathbb{C}$ be $\hat{\mu}$ for some μ on E ?
- If $\hat{\mu} = \hat{\nu}$ does $\mu = \nu$?

Definition 2.7. Let V be a real vector space. A function $f : V \rightarrow \mathbb{C}$ is of *positive type* if

- (i) for all $n \in \mathbb{N}$, if $\lambda_1, \dots, \lambda_n \in V$ then $(f(\lambda_i - \lambda_j))_{i,j=1}^n$ is a positive semi-definite complex $n \times n$ matrix;
- (ii) f is continuous on all finite-dimensional subspaces of V .

Definition 2.8. A matrix A is *positive semi-definite* if $A^\top = \bar{A}$ and $\langle A\xi, \xi \rangle_{\mathbb{C}^n} \geq 0$.

Proposition 2.9. For μ a probability measure on a separable Banach space E , $\hat{\mu} : E^* \rightarrow \mathbb{C}$ is of positive type with $\hat{\mu}(0) = 1$ and is continuous on E^* .

Proof. First observe that $\hat{\mu}(0) = \int_E 1 d\mu = \mu(E) = 1$. If $\lambda_1, \dots, \lambda_n \in E^*$ and $\xi_1, \dots, \xi_n \in \mathbb{C}$ then

$$\sum_{k,j=1}^n \hat{\mu}(\lambda_k - \lambda_j) \xi_k \bar{\xi}_j = \int_E \left| \sum_{j=1}^n e^{i\lambda_j(t)} \xi_j \right|^2 d\mu(t) \geq 0$$

and is clearly Hermitian since

$$\begin{aligned}\hat{\mu}(\lambda_j - \lambda_k) &= \int_E e^{i(\lambda_j(x) - \lambda_k(x))} d\mu(x) \\ &= \int_E e^{-i(\lambda_k(x) - \lambda_j(x))} d\mu(x) \\ &= \hat{\mu}(\lambda_k - \lambda_j).\end{aligned}$$

As for continuity, prove this as an exercise using the Dominated Convergence Theorem. \square

Theorem 2.10. (Bochner's Theorem. [RS]) *For a finite-dimensional vector space V (with the usual topology), the set of Fourier transforms of probability measures on V is precisely the set of $k : F^* \rightarrow \mathbb{C}$ of positive type with $k(0) = 1$. Moreover, each such k determines a unique probability measure μ , so $\hat{\mu} = \hat{\nu} \iff \mu = \nu$ on finite-dimensional spaces.*

3 Gaussian Measures on Finite-Dimensional Spaces

3.1 Gaussian Measures

Recall that we have standard Gaussian measure γ^n on \mathbb{R}^n :

$$\gamma^n(A) := (2\pi)^{-n/2} \int_A e^{-\|x\|^2/2} dx.$$

This is a probability measure, since

$$\begin{aligned} \int_{\mathbb{R}^n} e^{-\|x\|^2/2} dx &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-(x_1^2 + \dots + x_n^2)/2} dx_1 \dots dx_n \\ &= \prod_{j=1}^n \int_{-\infty}^{\infty} e^{-x_j^2/2} dx_j \end{aligned}$$

Also

$$\begin{aligned} \left(\int_{-\infty}^{\infty} e^{-x^2/2} dx \right)^2 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} dx dy \\ &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2/2} r dr d\theta \\ &= 2\pi [-e^{-r^2/2}]_0^{2\pi} \\ &= 2\pi \end{aligned}$$

Lemma 3.1. For C a positive definite matrix and A a symmetric $n \times n$ matrix,

(i) $\int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle Cx, x \rangle} dx = (2\pi)^{n/2} (\det C)^{-1/2};$

(ii) $\text{tr}(AC^{-1}) = \frac{\sqrt{\det C}}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \langle Ax, x \rangle e^{-\frac{1}{2}\langle Cx, x \rangle} dx;$

(iii) $(C^{-1})_{ij} = \frac{\sqrt{\det C}}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} x_i x_j e^{-\frac{1}{2}\langle Cx, x \rangle} dx.$

Note. C positive definite $\implies \langle Cx, x \rangle \geq \lambda \|x\|^2$ for all x , where λ is the smallest eigenvalue of C . So $e^{-\frac{1}{2}\langle Cx, x \rangle} \leq e^{-\frac{1}{2}\lambda \|x\|^2}$ and so all the above integrals exist.

Proof. (i) Diagonalize C as $C = U^{-1}\Lambda U$ with U orthogonal ($U^* = U^{-1}$) and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, $\lambda_j > 0$.

$$\begin{aligned} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle Cx, x \rangle} dx &= \int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle \Lambda Ux, Ux \rangle} dx \\ &= \int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle \Lambda y, y \rangle} dy \text{ with } y := Ux, U_* \lambda^n = \lambda^n \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\lambda_1 x_1^2 + \dots + \lambda_n x_n^2)} dx_1 \dots dx_n \\ &= \prod_{j=1}^n \int_{-\infty}^{\infty} e^{-\frac{1}{2}\lambda_j x_j^2} dx_j \\ &= \prod_{j=1}^n \int_{-\infty}^{\infty} e^{-\frac{1}{2}y_j^2} \frac{dy_j}{\lambda_j^{1/2}} \\ &= (2\pi)^{n/2} (\det C)^{-1/2} \end{aligned}$$

(ii) Take $h > 0$ so small that $C + hA$ is positive definite. By (i),

$$(2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle(C+hA)x,x\rangle} dx = (\det(C + hA))^{-1/2}.$$

Now take $\frac{d}{dh}\Big|_{h=0}$:

$$\begin{aligned} (2\pi)^{-n/2} \int_{\mathbb{R}^n} -\frac{1}{2}\langle Ax, x\rangle e^{-\frac{1}{2}\langle Cx,x\rangle} dx &= \frac{d}{dh}\Big|_{h=0} (\det(I + hAC^{-1}) \det C)^{-1/2} \\ &= -\frac{1}{2}(\det C)^{-1/2} \operatorname{tr}(AC^{-1}) \end{aligned}$$

since $\frac{d}{dh}\Big|_{h=0} \det(I + hK) = \operatorname{tr} K$ for any matrix K .

(iii) Apply (ii) with $A_{pq} = 0$ unless $(p, q) = (i, j)$ or (j, i) , otherwise $A_{ij} = A_{ji} = 1$, so $\operatorname{tr} AB = B_{ji} + B_{ij}$, for $B = C^{-1}$, and $\langle Ax, x\rangle = x_i x_j + x_j x_i$. \square

Remark 3.2. If $\{V, \langle \cdot, \cdot \rangle^\sim\}$ is an n -dimensional inner product space then $\langle \cdot, \cdot \rangle^\sim$ determines a ‘‘Lebesgue measure’’ $\lambda^{\langle \cdot, \cdot \rangle^\sim}$ on V . For this take an isometry $u : \mathbb{R}^n \rightarrow V$ with $\langle u(x), u(y) \rangle^\sim = \langle x, y \rangle$, so $u(x) = x_1 e_1 + \dots + x_n e_n$ for some orthonormal basis e_1, \dots, e_n of V . Set $\lambda^{\langle \cdot, \cdot \rangle^\sim} := u_* \lambda^n$. (So the ‘‘unit cube’’ spanned by e_1, \dots, e_n has $\lambda^{\langle \cdot, \cdot \rangle^\sim}$ -measure 1.) As an exercise, check that this does not depend on the choice of u .

Example 3.3. $V = \mathbb{R}^n$ with $\langle x, y \rangle^\sim = \langle Cx, y \rangle$ for some positive definite C . Write $C = U^{-1}\Lambda U$ with U orthogonal, $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Then $\sqrt{C} = U^{-1}\Lambda^{1/2}U$, where $\Lambda^{1/2} := \operatorname{diag}(\lambda_1^{1/2}, \dots, \lambda_n^{1/2})$ (the unique positive definite matrix K such that $KK = C$). $(\sqrt{C})^* = \sqrt{C}$ and $\sqrt{C}\sqrt{C} = C$. Define $u : \mathbb{R}^n \rightarrow V$ by $u(x) := (\sqrt{C})^{-1}x$, so $\langle u(x), u(y) \rangle^\sim = \langle x, y \rangle$. By definition, $\lambda^{\langle \cdot, \cdot \rangle^\sim} := u_* \lambda^n$.

$$\begin{aligned} \int_V f d\lambda^{\langle \cdot, \cdot \rangle^\sim} &= \int_{\mathbb{R}^n} f(u(x)) d\lambda^n(x) \\ &= \int_{\mathbb{R}^n} f(\sqrt{C}^{-1}x) d\lambda^n(x) \\ &= \det \sqrt{C} \int_{\mathbb{R}^n} f(y) dy. \end{aligned}$$

So $\lambda^{\langle C, \cdot \rangle} = \det \sqrt{C} \lambda^n = (\det C)^{1/2} \lambda^n$.

Definition 3.4. Let $\{V, \langle \cdot, \cdot \rangle^\sim\}$ be a finite-dimensional inner product space. The *standard Gaussian measure* $\gamma^{\langle \cdot, \cdot \rangle^\sim}$ on V is

$$\gamma^{\langle \cdot, \cdot \rangle^\sim}(A) := (2\pi)^{-n/2} \int_A e^{-\langle x, x \rangle^\sim / 2} d\lambda^{\langle \cdot, \cdot \rangle^\sim}(x)$$

for $A \in \mathcal{B}(V)$.

Remarks 3.5. (i) If $\dim V = n$ and $u : \mathbb{R}^n \rightarrow V$ is an isometry, then $\gamma^{\langle \cdot, \cdot \rangle^\sim} = u_*(\gamma^n)$, so $\gamma^{\langle \cdot, \cdot \rangle^\sim}$ is a probability measure.

(ii) If $V = \mathbb{R}^n$ with $\langle x, y \rangle^\sim := \langle Cx, y \rangle$, C as before, then

$$\gamma^{\langle \cdot, \cdot \rangle^\sim}(A) = (2\pi)^{-n/2} (\det C)^{1/2} \int_A e^{-\frac{1}{2}\langle Cx, x \rangle} dx.$$

Definitions 3.6. For V a finite-dimensional real vector space a (*centred*) *Gaussian measure* on V is one of the form $\mu = T_*\gamma^n$ for some $T \in \mathbb{L}(\mathbb{R}^n; V)$. It is *non-degenerate* if T is surjective.

Remark 3.7. In general, Gaussian measures may not be “centred”. They include $\mu = A_*\gamma^n$ for A affine.

Remark 3.8. A Gaussian measure μ is non-degenerate iff it is strictly positive (i) and iff it is $\gamma^{\langle \cdot, \cdot \rangle^\sim}$ for some $\langle \cdot, \cdot \rangle^\sim$ on V (ii). If $H, \langle \cdot, \cdot \rangle_H$ is a Hilbert space and $T : H \rightarrow V$ is linear and surjective, we get $\langle \cdot, \cdot \rangle_T$ on V by $\langle u, v \rangle_T := \langle \tilde{T}^{-1}(u), \tilde{T}^{-1}(v) \rangle_H$, where $\tilde{T} := T|_{(\ker T)^\perp} : (\ker T)^\perp \rightarrow V$ is bijective. This way we get a “quotient inner product”. For (ii), take $\langle \cdot, \cdot \rangle^\sim = \langle \cdot, \cdot \rangle_T$. For (i) note that $T_*\gamma^n(A) = 0$ if $A \cap T(\mathbb{R}^n) = \emptyset$. If T is not surjective, $T(\mathbb{R}^n)$ is a subspace not equal to V , so there exists open balls that do not intersect $T(\mathbb{R}^n)$, which contradicts strict positivity.

3.2 Fourier Transforms of Gaussian Measures

Lemma 3.9. For $\alpha \in \mathbb{C}$, $y \in \mathbb{R}^n$, and C positive-definite,

$$(2\pi)^{-n/2}(\det C)^{1/2} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle Cx, x \rangle} e^{\alpha\langle x, y \rangle} dx = e^{\frac{1}{2}\alpha^2\langle C^{-1}y, y \rangle}.$$

Proof. First consider $\alpha \in \mathbb{R}$:

$$\begin{aligned} \text{LHS} &= (2\pi)^{-n/2}(\det C)^{1/2} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle C(x-\alpha C^{-1}y), x-\alpha C^{-1}y \rangle} e^{\frac{1}{2}\alpha\langle y, C^{-1}y \rangle} dx \\ &= (2\pi)^{-n/2}(\det C)^{1/2} e^{\frac{1}{2}\alpha^2\langle y, C^{-1}y \rangle} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle Cx', x' \rangle} dx' \\ &= e^{\frac{1}{2}\alpha^2\langle y, C^{-1}y \rangle} \text{ by Lemma 3.1.} \end{aligned}$$

For $\alpha \in \mathbb{C}$, note that the LHS and RHS are holomorphic in α on the whole of \mathbb{C} and agree on \mathbb{R} , so they agree on \mathbb{C} . \square

Corollary 3.10. The Fourier transform of the standard Gaussian measure on \mathbb{R}^n satisfies

$$\widehat{\gamma^n}(\ell) = \exp\left(-\frac{1}{2}\|\ell^\sharp\|_{\mathbb{R}^n}^2\right) = \exp\left(-\frac{1}{2}\|\ell\|_{(\mathbb{R}^n)^*}^2\right).$$

Recall that if $\{V, \langle \cdot, \cdot \rangle_V\}$ is a real Hilbert space with $\dim V \leq \infty$, then V^* has a natural inner product making it a Hilbert space, with Riesz isometric isomorphism $V^* \rightarrow V : \ell \mapsto \ell^\sharp$ so that $\ell(x) = \langle \ell^\sharp, x \rangle_V$ for all $x \in V$. If $\{e_j\}_{j=1}^\infty$ is an orthonormal basis, then $\langle \ell^1, \ell^2 \rangle_{V^*} = \sum_{j=1}^\infty \ell^1(e_j)\ell^2(e_j)$, since $V^{**} \cong V$ canonically, so an inner product on V^* gives one on V .

Lemma 3.11. If E, F are Banach spaces with $T \in \mathbb{L}(E; F)$ surjective then $T^* \in \mathbb{L}(F^*; E^*)$ is injective.

Proof. If $T^*(\ell) = 0$, then $T^*(\ell)(x) = \ell(T(x)) = 0$ for all $x \in E$, so $\ell = 0$ since T is surjective. \square

Proposition 3.12. A probability measure μ on a finite-dimensional vector space V is non-degenerate Gaussian if, and only if, $\hat{\mu}(\ell) = e^{-\frac{1}{2}(\|\ell\|)^2}$ for all $\ell \in V^*$ for some inner product $\langle \cdot, \cdot \rangle$ on V^* .

Proof. (\implies) If $\mu = T_*\gamma^n$ for some surjective $T : \mathbb{R}^n \rightarrow V$, then, for $\ell \in V^*$,

$$\hat{\mu}(\ell) = \widehat{T_*\gamma^n}(\ell) = \widehat{\gamma^n}(T^*(\ell)),$$

by Remark 2.6. Moreover,

$$\widehat{\gamma^n}(T^*(\ell)) = e^{-\frac{1}{2}\|T^*(\ell)\|_{(\mathbb{R}^n)^*}^2} = e^{-\frac{1}{2}(\|\ell\|')^2},$$

where $\langle \ell^1, \ell^2 \rangle' := \langle T^*(\ell^1), T^*(\ell^2) \rangle_{(\mathbb{R}^n)^*}$, an inner product by Lemma 3.11.

(\impliedby) If $\hat{\mu}(\ell) = e^{-\frac{1}{2}(\|\ell\|')^2}$ for some $\langle \cdot, \cdot \rangle'$ on V^* , take $T : \mathbb{R}^m \rightarrow \{V, \langle \cdot, \cdot \rangle'\} \sim$ an isometry, where $m = \dim V$ and $\langle \cdot, \cdot \rangle \sim$ is the inner product on V corresponding to $\langle \cdot, \cdot \rangle'$ on V^* . Then, for all $\ell \in V^*$,

$$\widehat{T_*\gamma^m}(\ell) = \widehat{\gamma^m}(T^*(\ell)) = e^{-\frac{1}{2}\|T^*(\ell)\|_{(\mathbb{R}^m)^*}^2} = \hat{\mu}(\ell).$$

Bochner's Theorem then implies that $\mu = T_*\gamma^m$. \square

Theorem 3.13. *A strictly positive measure μ on a finite-dimensional vector space V is Gaussian if and only if $\ell_*\mu$ is a non-degenerate Gaussian measure on \mathbb{R} for all $\ell \in V^* \setminus \{0\}$. If so, $\ell \in L^2(V, \mu; \mathbb{R})$ for all $\ell \in V^*$ and $\hat{\mu}(\ell) = e^{-\frac{1}{2}\|\ell\|_{L^2}^2}$.*

Proof. (i) μ is non-degenerate Gaussian $\implies \mu = T_*\gamma^n$ for some surjective $T \in \mathbb{L}(\mathbb{R}^n; V) \implies \ell_*\mu = \ell_*T_*\gamma^n = (\ell \circ T)_*\gamma^n$ is non-degenerate Gaussian on \mathbb{R} , since $\ell \circ T$ is onto if $\ell \neq 0$.

(ii) Suppose that $\ell \in V^*$ is non-zero, so $\ell_*\mu$ is non-degenerate Gaussian on \mathbb{R} . Then $\ell_*\mu = \gamma^{\langle \cdot, \cdot \rangle_\ell}$ for some $\langle \cdot, \cdot \rangle_\ell$ on \mathbb{R} . Therefore, $\exists c(\ell) > 0$ such that

$$\ell_*\mu(A) = \int_A \frac{1}{\sqrt{2\pi}} e^{-c(\ell)t^2/2} c(\ell)^{1/2} dt$$

i.e., $\langle s, t \rangle_\ell = c(\ell)st$. Therefore, $\hat{\mu}(\ell) = \widehat{\ell_*\mu}(1) = e^{-1/(2c(\ell))}$ since $\widehat{\ell_*\mu}(s) = e^{-s^2/(2c(\ell))}$ by Corollary 3.10. Now

$$\begin{aligned} \|\ell\|_{L^2}^2 &= \int_V |\ell(x)|^2 d\mu(x) \\ &= \int_{\mathbb{R}} t^2 d(\ell_*\mu)(t) \\ &= c(\ell)^{1/2} \int_{\mathbb{R}} t^2 (2\pi)^{-1/2} e^{-c(\ell)t^2/2} dt \\ &= c(\ell)^{-1} \text{ by Lemma 3.1 (iii)} \\ &< \infty \end{aligned}$$

Therefore, $\hat{\mu}(\ell) = e^{-\frac{1}{2}\|\ell\|_{L^2}^2}$ as required. Next note that the quotient map $V^* \rightarrow L^2(V, \mu; \mathbb{R}) : \ell \mapsto [\ell]$ is injective since $\ell \in V^*$ and

$$\begin{aligned} \ell = 0 \text{ in } L^2 &\implies \ell(x) = 0 \text{ almost everywhere in } V \\ &\implies \ker \ell \text{ has full measure} \\ &\implies \mu \text{ not strictly positive unless } \ell \equiv 0 \end{aligned}$$

So now define $\langle \cdot, \cdot \rangle'$ on V^* by $\langle \ell^1, \ell^2 \rangle' := \langle [\ell^1], [\ell^2] \rangle_{L^2}$ and apply Proposition 3.12. \square

4 Gaussian Measures on Banach Spaces

Definitions 4.1. Let E be a separable Banach space. A Borel probability measure μ on E is said to be *Gaussian* if $\ell_*\mu$ is Gaussian on \mathbb{R} for all $\ell \in E^*$. Such a μ is *non-degenerate* if it is strictly positive.

Remark 4.2. By Theorem 3.13, this agrees with the finite-dimensional definition in the non-degenerate case; a slight modification of the proof of Theorem 3.13 handles the general case.

Lemma 4.3. *If μ on E is strictly positive and $\ell \in E^* \setminus \{0\}$ then $\ell_*\mu$ on \mathbb{R} is strictly positive.*

Proof. Take $U \subseteq \mathbb{R}$ non-empty and open. Then $\ell_*\mu(U) = \mu(\ell^{-1}(U)) > 0$ since $\ell^{-1}(U)$ is open and non-empty in E , since ℓ is continuous and onto. \square

Theorem 4.4. *If γ is Gaussian and non-degenerate on E then for all $\ell \in E^*$, $\ell \in L^2(E, \gamma; \mathbb{R})$ and $\hat{\gamma}(\ell) = e^{-\frac{1}{2}\|\ell\|_{L^2}^2}$.*

The proof of this mimics that of Theorem 3.13 and is omitted. However, we have not yet established whether or not there are any non-degenerate Gaussian measures on infinite-dimensional spaces!

5 Cylinder Set Measures

Definition 5.1. Let E be a separable Banach space and let

$$\mathcal{A}(E) := \{T \in \mathbb{L}(E; F) \mid \dim F < \infty, T \text{ onto}\}.$$

We will write F_T for F if $T \in \mathcal{A}(E), T \in \mathbb{L}(E; F)$. A *cylinder set measure* (or CSM) on E is a family $\{\mu_T\}_{T \in \mathcal{A}(E)}$ of probability measures μ_T on $F_T, T \in \mathcal{A}(E)$, such that if we have

$$\begin{array}{ccc} E & \xrightarrow{T} & F_T \\ & \searrow S & \downarrow \pi_{ST} \\ & & F_S \end{array}$$

then $\mu_S = (\pi_{ST})_*(\mu_T)$.

Examples 5.2. (i) If μ is a probability measure on E define $\mu_T := T_*(\mu)$ for each $T \in \mathcal{A}(E)$. Then if $\pi_{ST} \circ T = S$ as above,

$$\mu_S = S_*(\mu) = (\pi_{ST} \circ T)_*(\mu) = (\pi_{ST})_*(T_*(\mu)) = (\pi_{ST})_*(\mu_T).$$

If a CSM $\{\mu_T\}_{T \in \mathcal{A}(E)}$ on E corresponds to a measure in this way (i.e., there is a measure μ on E such that $\mu_T = T_*\mu$ on F_T) then we say that it “is” a measure, although we don’t yet know that it is unique.

(ii) If $\dim E < \infty$ every CSM on E is a measure — just take $\mu = \mu_{\text{id}}$ for $\text{id} : E \rightarrow E$ the identity map.

(iii) A real Hilbert space $\{H, \langle \cdot, \cdot \rangle_H\}$. We have a canonical Gaussian CSM on $H, \{\gamma_T^H \mid T \in \mathcal{A}(H)\}$. γ_T^H on F_T (where $T : H \rightarrow F_T$ is onto) is defined by $\gamma_T^H := \gamma^{\langle \cdot, \cdot \rangle_T}$, where $\langle \cdot, \cdot \rangle_T$ is the quotient inner product on F_T :

$$\langle u, v \rangle_T := \left\langle T|_{(\ker T)^\perp}^{-1} u, T|_{(\ker T)^\perp}^{-1} v \right\rangle_H.$$

Equivalently, $\langle \cdot, \cdot \rangle_T$ on F_T is determined by $\langle \cdot, \cdot \rangle^T$ on F_T^* , where

$$\langle \ell^1, \ell^2 \rangle^T = \langle T^* \ell^1, T^* \ell^2 \rangle_{H^*}, \quad (5.1)$$

so $\widehat{\gamma_T^H}(\ell) = \exp(-\frac{1}{2} \|T^* \ell\|_{H^*}^2)$.

Exercise 5.3. Show that $\{\gamma_T^H \mid T \in \mathcal{A}(H)\}$ is a CSM. Hint: Use Fourier transforms.

Proposition 5.4. For $T = \ell \in H^*, \ell \neq 0, \gamma_\ell^H$ on $\mathbb{R} = F_\ell$ is given by

$$\gamma_\ell^H(A) = \frac{1}{\sqrt{2\pi} \|\ell^\sharp\|_H} \int_A \exp\left(\frac{-t^2}{2\|\ell^\sharp\|_H^2}\right) dt$$

where $\ell^\sharp \in H$ is the Riesz representative of $\ell \in H^*$.

Proof. Method One. Use (5.1). $\widehat{\gamma}_\ell^H(s) = \exp(-\frac{1}{2}\|\ell^*s\|_{H^*}^2)$ for $s \in \mathbb{R} \cong \mathbb{R}^*$, where $\ell^* : \mathbb{R} \rightarrow H \cong H^*$ is $t \mapsto t\ell^\sharp$ since

$$\langle t\ell^\sharp, h \rangle_H = t\langle \ell^\sharp, h \rangle_H = t\ell(h) = \langle t, \ell(h) \rangle_{\mathbb{R}} = \langle \ell^*(t), h \rangle_H$$

as required. Therefore, $\gamma_\ell^H(s) = \exp(-\frac{1}{2}\|\ell^\sharp\|_H^2 s^2)$, which is the Fourier transform of the measure given.

Method Two. $(\ker \ell)^\perp = \{s\ell^\sharp | s \in \mathbb{R}\}$ since this is one-dimensional ($\ell \neq 0$) and is $h \in \ker \ell$ then

$$\langle h, s\ell^\sharp \rangle_H = s\langle h, \ell^\sharp \rangle_H = s\ell(h) = 0 \text{ for all } s.$$

So set $\tilde{\ell} = \ell|_{(\ker \ell)^\perp}$. $\tilde{\ell}$ is

$$s\ell^\sharp \mapsto \ell(s\ell^\sharp) = s\ell(\ell^\sharp) = s\|\ell^\sharp\|_H^2.$$

$\tilde{\ell}^{-1} : \mathbb{R} \rightarrow (\ker \ell)^\perp \subseteq H$ is given by $t \mapsto \frac{t\ell^\sharp}{\|\ell^\sharp\|_H^2}$, therefore

$$\langle s, t \rangle_\ell = \left\langle \frac{s\ell^\sharp}{\|\ell^\sharp\|_H^2}, \frac{t\ell^\sharp}{\|\ell^\sharp\|_H^2} \right\rangle_H = \frac{st}{\|\ell^\sharp\|_H^2},$$

so the measure is as given. □

Definition 5.5. The *Fourier transform* of a CSM $\{\mu_T | T \in \mathcal{A}(E)\}$ is defined to be $\widehat{\mu} : E^* \rightarrow \mathbb{C}$ given by

$$\begin{aligned} \widehat{\mu}(\ell) &:= \widehat{\mu}_\ell(1) \\ &= \int_{\mathbb{R}} e^{-it} d\mu_\ell(t) \text{ (the Fourier transform of the measure } \mu_\ell \text{ on } \mathbb{R}) \end{aligned}$$

for $\ell \neq 0$, and $\widehat{\mu}(0) := 1$. This agrees with the definition when the CSM μ is a measure, by Remark 2.5.

Proposition 5.6. For the canonical Gaussian CSM on H , $\{\gamma_T^H\}_{T \in \mathcal{A}(E)}$, if $\ell \in H^*$ is non-zero,

$$(i) \quad \widehat{\gamma}_\ell^H(\ell) = \exp(-\frac{1}{2}\|\ell^\sharp\|_H^2);$$

$$(ii) \quad \int_{\mathbb{R}} t^2 d\gamma_\ell^H(t) = \|\ell^\sharp\|_H^2.$$

Definition 5.7. Suppose that $\theta : E_1 \rightarrow E_2$ is a linear map of separable Banach spaces. Given a CSM $\{\mu_T | T \in \mathcal{A}(E_1)\}$ on E_1 we get a *push-forward* CSM $\{\theta_*(\mu.)_S | S \in \mathcal{A}(E_2)\}$ on E_2 by

$$\theta_*(\mu.)_S = \mu_{S \circ \theta}$$

if $S \circ \theta$ is onto. If $S \circ \theta$ is not onto let \tilde{F} be the image of $S \circ \theta$, $i : \tilde{F} \hookrightarrow F_S$ the inclusion, and define

$$\theta_*(\mu.)_S = i_*(\widetilde{\mu_{S \circ \theta}})$$

where $\widetilde{S \circ \theta} : E_1 \rightarrow \tilde{F}$ is such that $i \circ \widetilde{S \circ \theta} = S \circ \theta$.

Definition 5.8. (i) Given a CSM $\{\mu_T\}_T$ on E and $\theta \in \mathbb{L}(E; G)$, G a separable Banach space, we say that θ *radonifies* $\{\mu_T\}_T$ if $\theta_*(\mu.)$ is a measure on G .

- (ii) $\theta \in \mathbb{L}(H; G)$, H a separable Hilbert space, G a separable Banach space, is γ -radonifying if $\{\theta_*(\gamma^H)_T\}_T$ is a measure on G , i.e. θ radonifies the canonical Gaussian CSM $\{\gamma_T^H\}_{T \in \mathcal{A}(H)}$ on H .

Examples 5.9. (i) If θ has finite rank then θ radonifies all CSMs. For example, $\theta_*(\mu) = \mu_\theta$ if $\theta : E \rightarrow G$ is onto and $\dim G < \infty$.

- (ii) If $\text{id} : E \rightarrow E$ is the identity then id radonifies $\{\mu_T\}_T$ if, and only if, $\{\mu_T\}_T$ is a measure.

Definitions 5.10. For H a separable Hilbert space and E a separable Banach space, if $i : H \rightarrow E$ is a continuous linear injective map with dense range that γ -radonifies we say that $i : H \rightarrow E$ is an *abstract Wiener space* (or AWS). For example, $L^2 \rightarrow L^1$. The measure induced on E is called the *abstract Wiener measure* of $i : H \rightarrow E$.

Proposition 5.11. *An abstract Wiener measure is a Gaussian measure.*

Proof. We need to show that $\ell_*\gamma$ is Gaussian on \mathbb{R} for all $\ell \in E^*$:

$$\begin{array}{ccc} H & \xrightarrow{i} & E, \gamma \\ & \searrow \ell \circ i & \downarrow \ell \\ & & \mathbb{R}, \ell_*(\gamma) \end{array}$$

$$\begin{aligned} \ell_*(\gamma) &= \ell_*(i_*(\gamma^H)) \\ &= (i_*(\gamma^H))_\ell \\ &= \gamma_{\ell \circ i}^H \end{aligned}$$

which is Gaussian if $\ell \neq 0$; if $\ell = 0$ we get δ_0 . □

Example 5.12. Classical Wiener space. Let

$$\begin{aligned} H &:= L_0^{2,1}([0, T]; \mathbb{R}^n) \\ &= \{\text{paths beginning at 0 with first derivative} \in L^2\} \\ &= \left\{ \sigma : [0, T] \rightarrow \mathbb{R}^n \mid \exists \phi \in L^2([0, T]; \mathbb{R}^n) \text{ with } \sigma(t) = \int_0^t \phi(s) ds \right\} \end{aligned}$$

So $\dot{\sigma}(s) = \phi(s)$ for almost all $s \in [0, T]$ and $\sigma(0) = 0$.

$$\langle \sigma^1, \sigma^2 \rangle_{L_0^{2,1}} = \int_0^T \langle \dot{\sigma}^1(s), \dot{\sigma}^2(s) \rangle_{\mathbb{R}^n} ds$$

The operator $\frac{d}{dt} : L_0^{2,1} \rightarrow L^2$ is an isometry of Hilbert spaces. Let

$$\begin{aligned} E &:= C_0([0, T]; \mathbb{R}^n) \\ &= \{\sigma : [0, T] \rightarrow \mathbb{R}^n \mid \sigma \text{ is continuous and } \sigma(0) = 0\} \end{aligned}$$

$$\|\sigma\|_E := \|\sigma\|_\infty := \sup_{0 \leq t \leq T} \|\sigma(t)\|_{\mathbb{R}^n}$$

Then the inclusion $i : H \hookrightarrow E$ is continuous and linear. By Cauchy-Schwarz, it is injective. The image is dense in E by the standard approximation theorems — e.g., polynomials p with $p(0) = 0$ are dense in C_0 (the Stone-Weierstrass Theorem).

Theorem 5.13. (Wiener, Gross et. al.) *The inclusion $i : L_0^{2,1} \hookrightarrow C_0$ is γ -radonifying.*

Definitions 5.14. The Gaussian measure γ induced on C_0 is *classical Wiener measure*. Also, C_0 , or $i : L_0^{2,1} \hookrightarrow C_0$, is called *classical Wiener space*. $L_0^{2,1}$ is called the corresponding *Cameron-Martin space* or the *reproducing kernel Hilbert space*.

Some questions to deal with:

(i) Is the map

$$\begin{aligned} &\text{probability measures on } E \rightarrow \text{CSMs on } E \\ &\mu \mapsto \{T_*(\mu) | T \in \mathcal{A}(E)\} \end{aligned}$$

injective?

(ii) How about Fourier transforms of measures in infinite dimensions? Do we have an analogue of Bochner's Theorem?

Lemma 5.15. *If E, G are separable Banach spaces, $\theta \in \mathbb{L}(E; G)$ and $\{\mu_T\}_{T \in \mathcal{A}(E)}$ is a CSM on E , then*

$$\widehat{\theta_*(\mu)}(\ell) = \widehat{\mu}(\theta^*(\ell))$$

for all $\ell \in G^*$. In particular, if $T \in \mathcal{A}(E)$, then $\widehat{\mu_T}(\ell) = \widehat{\mu}(T^*(\ell))$ for all $\ell \in F_T^*$.

Proof. If $\ell \neq 0$ then

$$\begin{aligned} \widehat{\theta_*(\mu)}(\ell) &= \widehat{\theta_*(\mu)}_\ell(1) \text{ by definition of } \widehat{\mu} \\ &= \widehat{\mu_{\ell \circ \theta}}(1) \text{ if } \ell \circ \theta \neq 0 \text{ by definition of } \theta_*(\mu) \\ &= \widehat{\mu_{\theta^*(\ell)}}(1) \\ &= \widehat{\mu}(\theta^*(\ell)) \text{ by definition of } \widehat{\mu}. \end{aligned}$$

If $\ell \circ \theta = 0$ then $\theta^*(\ell) = 0$ so RHS = 1 (probability measure) but LHS = 1 since $\theta_*(\mu) = \delta_0$. \square

Theorem 5.16. (Extended Bochner Theorem.) *The functions of positive type $f : E^* \rightarrow \mathbb{C}$ with $f(0) = 1$ are precisely the Fourier transforms of CSMs on E and $\widehat{\mu} = \widehat{\nu} \implies \{\mu_T\}_T = \{\nu_T\}_T$.*

Proof. Given $f : E^* \rightarrow \mathbb{C}$ of positive type and $T \in \mathcal{A}(E)$ (so that $T : E \rightarrow F_T$ is surjective), the composition $f \circ T^* : F_T^* \rightarrow \mathbb{C}$ is continuous, since $\dim F_T^* < \infty$, and positive, and so is of positive type. Therefore, by the finite-dimensional Bochner Theorem, we get μ_T on F_T with $\widehat{\mu_T} = f \circ T^*$. One can check that $\{\mu_T\}_{T \in \mathcal{A}(E)}$ forms a CSM.

This argument shows that $\widehat{\mu}$ determines $\{\mu_T\}_{T \in \mathcal{A}(E)}$.

Given a CSM $\{\mu_T\}_{T \in \mathcal{A}(E)}$ on E and $\ell_1, \dots, \ell_n \in E^*$, we need to show that

(a) $\forall \xi_1, \dots, \xi_n \in \mathbb{C}, \sum_{i,j=1}^n \widehat{\mu}(\ell_i - \ell_j) \xi_i \overline{\xi_j} \geq 0;$

(b) if $F = \text{span}\{\ell_1, \dots, \ell_n\}$, then $\widehat{\mu}$ is continuous on F .

(a) Let $\tilde{\ell}_1, \dots, \tilde{\ell}_N$ be a basis for F . Define $T : E \rightarrow \mathbb{R}^N$ by $T(x) := (\tilde{\ell}_1(x), \dots, \tilde{\ell}_N(x))$, which is surjective. Therefore, we get μ_T on \mathbb{R}^N . Also, since T is onto, $T^* : (\mathbb{R}^N)^* \rightarrow E^*$ is injective. Its

image is F since it sends the dual basis in $(\mathbb{R}^N)^*$ to $\{\tilde{\ell}_j\}_{j=1}^N$. Take $e'_j \in (\mathbb{R}^N)^*$ such that $T^*(e'_j) = \ell_j$. Then

$$\begin{aligned} \sum_{i,j} \widehat{\mu}(\ell_i - \ell_j) \xi_i \bar{\xi}_j &= \sum_{i,j} \widehat{\mu} \cdot T^*(e'_i - e'_j) \xi_i \bar{\xi}_j \\ &= \sum_{i,j} \widehat{\mu}_T(e'_i - e'_j) \xi_i \bar{\xi}_j \text{ by Lemma 5.15} \\ &\geq 0 \text{ since } \mu_T \text{ is a measure} \end{aligned}$$

(b) $\widehat{\mu}_T = \widehat{\mu} \circ T^*$ by Lemma 5.15, therefore $\widehat{\mu}|_F = \widehat{\mu}_T \circ (T^*|_F^{-1})$, which is continuous since $\widehat{\mu}_T$ is. \square

Theorem 5.17. *Let E be a separable Banach space with finite measures μ, ν on E . Then*

(i) *if $T_*\mu = T_*\nu$ for all $T \in \mathcal{A}(E)$ then $\mu = \nu$;*

(ii) *if $\widehat{\mu} = \widehat{\nu}$ then $\mu = \nu$.*

Proof. By the Extended Bochner Theorem, Theorem 5.16, (i) \implies (ii), so we need only prove the first part. We define the *cylinder sets* $\text{Cyl}(E) := \{T^{-1}(B) | B \in \mathcal{B}(F_T), T \in \mathcal{A}(E)\}$. This is an algebra of subsets, but not a σ -algebra if $\dim E = \infty$. Given a probability measure μ on E and $T \in \mathcal{A}(E)$, $A = T^{-1}(B)$ for some $B \in \mathcal{B}(F_T)$, $\mu(A) = T_*(\mu)(B) = \mu_T(B)$. Therefore, the CSM $\{T_*\mu | T \in \mathcal{A}(E)\}$ determines $\mu(A)$ for all $A \in \text{Cyl}(E)$. Thus, the theorem follows from the following Lemma 5.18: \square

Lemma 5.18. *If E is a separable Banach space then $\mathcal{B}(E) = \sigma(\text{Cyl}(E))$, the smallest σ -algebra containing $\text{Cyl}(E)$.*

Theorem 5.19. (Uniqueness of Carathéodory's Extension. [RW1].) *Let μ, ν be finite measures on a measurable space $\{X, \mathcal{A}\}$ and let $\mathcal{A}^0 \subset \mathcal{A}$ be an algebra of subsets of X such that $\sigma(\mathcal{A}^0) = \mathcal{A}$. Then if $\mu = \nu$ on \mathcal{A}^0 , $\mu = \nu$ on \mathcal{A} as well. (This actually holds if \mathcal{A}^0 is just a π -system, one that is closed under finite intersections.)*

Proof of Lemma 5.18. Since $T : E \rightarrow F_T$ is continuous it is measurable, and $T^{-1}(B)$ is Borel if B is Borel, so $\text{Cyl}(E) \subseteq \mathcal{B}(E)$.

Consider the special case that $E \subseteq C([0, T]; \mathbb{R})$ is a closed subspace. Then $\mathcal{B}(E) = \{E \cap U | U \in \mathcal{B}(C([0, T]; \mathbb{R}))\}$ since

- the RHS is a σ -algebra;
- $\text{RHS} \subseteq \mathcal{B}(E)$ since the inclusion $i : E \hookrightarrow C([0, T]; \mathbb{R})$ is continuous, therefore measurable;
- all open balls in E lie in the RHS.

Take $x_0 \in E$ and $\varepsilon > 0$. We show that $\overline{B_\varepsilon(\sigma_0)} \in \sigma(\text{Cyl}(E))$. Since $\mathcal{B}(E)$ is generated by all such balls, the result will follow. For this, let $\{q_1, q_2, \dots\}$ be an enumeration of $\mathbb{Q} \cap [0, T]$. So

$$\begin{aligned} \overline{B_\varepsilon(x_0)} &= \{x \in E | \forall r \in [0, 1], |x(r) - x_0(r)| \leq \varepsilon\} \\ &= \{x \in E | |x(q_i) - x_0(q_i)| \leq \varepsilon, 1 \leq i < \infty\} \\ &= \bigcap_{i=1}^{\infty} \{x \in E | |x(q_i) - x_0(q_i)| \leq \varepsilon\} \\ &\in \text{Cyl}(E), \end{aligned}$$

because

$$\{x \in E \mid |x(q_i) - x_0(q_i)| \leq \varepsilon\} = \text{ev}_{q_i}^{-1} \left(\overline{B_\varepsilon^{\mathbb{R}}(x_0(q_i))} \right) \in \text{Cyl}(E).$$

For the general case we use

Theorem 5.20. (Banach-Mazur. [BP]) *Any separable Banach space is isometrically isomorphic to a closed subspace of $C([0, T]; \mathbb{R})$.*

Such an isomorphism maps $E \rightarrow \tilde{E} \subseteq C([0, T]; \mathbb{R})$; it maps $\text{Cyl}(E)$ to $\text{Cyl}(\tilde{E})$ and $\mathcal{B}(E)$ and $\mathcal{B}(\tilde{E})$ bijectively. We proved the result for \tilde{E} , so it is true for E . \square

Remark 5.21. The proof showed that $\mathcal{B}(E) = \sigma\{\ell \mid \ell \in E^*\}$ = smallest σ -algebra such that each $\ell \in E^*$ is measurable as a function $\ell : E \rightarrow \mathbb{R}$, and that for E closed in $C([0, T]; \mathbb{R})$, $\mathcal{B}(E) = \sigma\{\text{ev}_q \mid q \in \mathbb{Q} \cap [0, 1]\}$, where $\text{ev}_q(x) := x(q)$ is the evaluation map.

6 The Paley-Wiener Map and the Structure of Gaussian Measures

6.1 Construction of the Paley-Wiener Integral

Let $i : H \rightarrow E$ be an AWS with measure γ . Let $j : E^* \rightarrow H \cong H^*$ be the adjoint of i , defined by $\langle j(\ell), h \rangle_H = \ell(i(h))$ for $h \in H$, i.e. $j(\ell) = (\ell \circ i)^\# = (i^*(\ell))^\#$. So $E^* \xrightarrow{j} H \xrightarrow{i} E$.

Lemma 6.1. (i) $j : E^* \rightarrow H$ is injective.

(ii) j has dense range (i.e. $\overline{j(E^*)} = H$).

Proof. (i)

$$\begin{aligned} j(\ell) = 0 &\implies (\ell \circ i)^\# = 0 \\ &\implies \ell \circ i = 0 \\ &\implies \ell|_{i(H)} = 0 \\ &\implies \ell = 0 \end{aligned}$$

since $i(H)$ is dense in E and ℓ is continuous.

(ii) Suppose that $h \perp j(E^*)$, i.e. $\langle h, j(\ell) \rangle_H = 0$ for all $\ell \in E^*$. Then $\ell(i(h)) = 0$ for all $\ell \in E^*$. So $i(h) = 0$ by the Hahn-Banach Theorem. So $h = 0$, since i is injective. So $j(E^*)$ is dense in H . \square

Lemma 6.2. Given Banach spaces F and G , a dense subspace $F_0 \subseteq F$, and a map $\alpha \in \mathbb{L}(F_0; G)$ such that $\exists k$ such that $\|\alpha(x)\|_G \leq k\|x\|_F$ for all $x \in F_0$, then there exists a unique $\tilde{\alpha} \in \mathbb{L}(F; G)$ such that $\tilde{\alpha}|_{F_0} = \alpha$. Also, $\|\tilde{\alpha}\| \leq k$. Moreover, if $\|\alpha(x)\|_G = k\|x\|_F$ for all $x \in F_0$, then $\|\tilde{\alpha}(x)\|_G = k\|x\|_F$ for all $x \in F$, and so $\tilde{\alpha}$ is an isometry if $k = 1$.

Proof. Let $x \in F$. Take $(x_n)_{n=1}^\infty$ in F_0 with $x_n \rightarrow x$ in F . Then

$$\|\alpha(x_n) - \alpha(x_m)\|_G = \|\alpha(x_n - x_m)\|_G \leq k\|x_n - x_m\|_F$$

and so $(\alpha(x_n))_{n=1}^\infty$ is Cauchy in G , and so it converges in G . Set $\tilde{\alpha}(x) = \lim_{n \rightarrow \infty} \alpha(x_n)$. Check that this is independent of the choice of the $x_n \rightarrow x$. So we get $\tilde{\alpha} : F \rightarrow G$ extending α . Check that it is linear and unique. For the last part,

$$\begin{aligned} \|\tilde{\alpha}(x)\|_G &= \left\| \lim_{n \rightarrow \infty} \alpha(x_n) \right\|_G \\ &= \lim_{n \rightarrow \infty} \|\alpha(x_n)\|_G \\ &\leq \lim_{n \rightarrow \infty} k\|x_n\|_F \\ &= k\|x\|_F \end{aligned}$$

Therefore, $\|\tilde{\alpha}\| \leq k$ and $\tilde{\alpha}$ is continuous. If $\|\alpha(x_n)\|_G = k\|x_n\|_F$ for all n , the above argument shows that $\|\tilde{\alpha}(x)\|_G = k\|x\|_F$ for all $x \in F$. \square

Theorem 6.3. If $\ell \in E^*$ then $\ell \in L^2(E, \gamma; \mathbb{R})$ with $\|\ell\|_{L^2} = \|j(\ell)\|_H$. Consequently, there is a unique continuous linear $I : H \rightarrow L^2(E, \gamma; \mathbb{R})$, with $I(h) := \langle h, - \rangle_{\tilde{H}}$, such that

$$\begin{array}{ccc} H & \xrightarrow{I} & L^2(E, \gamma; \mathbb{R}) \\ & \swarrow j & \nearrow \pi: \ell \mapsto [\ell] \\ & E^* & \end{array}$$

Moreover, $\|I(h)\|_{L^2} = \|h\|_H$, so I is an isometry into $L^2(E, \gamma; \mathbb{R})$.

Proof. Let $\ell \in E^*$, $\ell \neq 0$.

$$\begin{aligned} \|\ell\|_{L^2}^2 &= \int_E \ell(x)^2 d\gamma(x) \\ &= \int_{\mathbb{R}} t^2 d(\ell_*(\gamma))(t) \\ &= \int_{\mathbb{R}} t^2 d\gamma_\ell^H(t) \\ &= \|(\ell \circ i)^\# \|_H^2 \text{ by Proposition 5.6 (ii)} \\ &= \|j(\ell)\|_H^2 < \infty \end{aligned}$$

For the ‘‘consequently’’ part, we apply Lemma 6.2 with $F_0 = j(E^*)$, $F = H$, $G = L^2(E, \gamma; \mathbb{R})$. \square

Definition 6.4. The isometry $I : H \rightarrow L^2(E, \gamma; \mathbb{R})$ is called the *Paley-Wiener map*. It is the unique extension to all of H of the natural map $j(E^*) \rightarrow L^2(E, \gamma; \mathbb{R})$ given by $j(\ell) \mapsto [\ell]_{L^2}$, which is well-defined by Lemma 6.1(i).

Remark 6.5. For $h \in H$, $I(h) = \lim_{n \rightarrow \infty} \ell_n$ in L^2 , where $\ell_n \in E^*$ with $j(\ell_n) \rightarrow h$ in H . We have $E^* \xrightarrow{j} H \xrightarrow{i} E$ with $\langle j(\ell), h \rangle_H = \ell(i(h))$, $j(\ell) = (\ell \circ i)^\# = (i(\ell))^\#$. $I : H \rightarrow L^2(E, \gamma; \mathbb{R})$ is isometric onto its image.

Remark 6.6. If $\dim H < \infty$ we can take $H = E$ and $i = \text{id}$, so $j : E^* \rightarrow H$ is $j(\ell) = (\ell \circ \text{id})^\# = \ell^\#$. In this case, j is the Riesz transform $H^* \rightarrow H$.

If $h \in H =$ the image of j , take ℓ_n such that $\ell_n^\# = h$ for all n , so $I(h) = \ell_n = \langle h, - \rangle_H$. Thus, in finite dimensions, $I(h) = \langle h, - \rangle_H$; thus, in infinite dimensions we sometimes write $\langle h, - \rangle_{\tilde{H}}$ for $I(h)$.

Note that $\langle h, x \rangle_H$ does not (in general) exist in the infinite-dimensional case. If $x \in E$, we can make classical sense of it if $x \in i(H)$, $x = i(k)$ for some $k \in H$: we use $\langle h, k \rangle_H$. If $h = \ell(j)$, $\ell \in E^*$, we use $\ell(x)$.

Now use $I(h) = \langle h, - \rangle_{\tilde{H}}$ — this is only defined as an element of $L^2(E, \gamma; \mathbb{R})$, so $I(h)(x)$ only makes sense up to sets of measure zero.

In classical Wiener space $C_0([0, T]; \mathbb{R}^n)$ with its Cameron-Martin space $H = L_0^{2,1}([0, T]; \mathbb{R}^n)$,

$$\begin{aligned} \langle h^1, h^2 \rangle_H &= \int_0^T \langle \dot{h}^1(s), \dot{h}^2(s) \rangle_{\mathbb{R}^n} ds \\ &= \text{‘‘} \int_0^T \langle \dot{h}^1(s), dh^2(s) \rangle_{\mathbb{R}^n} \text{’’ (Stieltjes)} \end{aligned}$$

‘‘ $dh(s)$ ’’ means ‘‘ $\dot{h}(s) ds$ ’’.

We often write $\langle h, - \rangle_H^\sim : C_0 \rightarrow \mathbb{R}$ as

$$\sigma \mapsto \int_0^T \langle \dot{h}(s), d\sigma(s) \rangle_{\mathbb{R}^n}.$$

This is only defined up to sets of Wiener measure zero, and is the *Paley-Wiener integral* of \dot{h} . However, this “line integral” exists even if the path σ is merely continuous; we do not need it be differentiable.

Definition 6.7. The *Paley-Wiener integral* for $f \in L^2([0, T]; \mathbb{R}^n)$ is

$$\left(\sigma \mapsto \int_0^T \langle f(s), d\sigma(s) \rangle_{\mathbb{R}^n} \right) := I \left(\int_0^T f(s) ds \right).$$

That is, take $h \in L_0^{2,1}([0, T]; \mathbb{R})$ such that $\dot{h} = f$. It is in $L^2(C_0, \gamma; \mathbb{R})$ as a function of σ .

$$H = L_0^{2,1} \begin{array}{c} \xrightarrow{\frac{d}{dt}} \\ \xleftarrow{\int_0^\cdot -ds} \end{array} L^2$$

Exercise 6.8. Let $\mu.$ be a CSM on E and let $\ell \in E^*$, $\ell \neq 0$. Prove that for $s \in \mathbb{R} \cong \mathbb{R}^*$, $\widehat{\mu}_\ell(s) = \widehat{\mu.}(s\ell)$.

Proposition 6.9. For any AWS $i : H \rightarrow E$, if $h \in H$, $h \neq 0$, then

$$I(h)_*(\gamma) = \gamma_{\langle h, - \rangle_H}^H.$$

Proof. If $h = j(\ell)$ for some $\ell \in E^*$, then $\ell \circ i = \langle h, - \rangle_H \in H^*$ and $I(h) = \ell$ by definition. Then $\ell_*(\gamma) = (\ell \circ i)_*(\gamma^H) = \gamma_{\langle h, - \rangle_H}^H$, as required.

In general, let $\ell_n \in E^*$ with $j(\ell_n) \rightarrow h$ in H , so $I(h) = \lim_{n \rightarrow \infty} [\ell_n]$ in L^2 . If $s \in \mathbb{R}$ then

$$\begin{aligned} \widehat{(\ell_n)_* \gamma}(s) &= \widehat{\gamma_{\ell_n \circ i}^H}(s) \\ &= e^{-\frac{1}{2}s^2 \|\ell_n\|_H^2} \text{ by Proposition 5.6 and Exercise 6.8} \\ &\rightarrow e^{-\frac{1}{2}s^2 \|h\|_H^2} \text{ as } n \rightarrow \infty \\ &= \widehat{\gamma_{\langle h, - \rangle}^H}(s) \text{ by Proposition 5.6 and Exercise 6.8.} \end{aligned}$$

But $\widehat{(\ell_n)_* \gamma}(s) = \widehat{\gamma}(s\ell_n) = \int_E e^{is\ell_n(x)} d\gamma(x)$. Now $\ell_n \rightarrow I(h)$ in L^2 and

$$\left| e^{is\ell_n(x)} - e^{isI(h)(x)} \right| \leq 2 \left| \sin \frac{s\ell_n(x) - sI(h)(x)}{2} \right|$$

since $e^{ix} - e^{iy} = 2ie^{i(x+y)/2} \sin \frac{x-y}{2}$

$$\begin{aligned} &\leq |s\ell_n(x) - sI(h)(x)| \\ &\rightarrow 0 \text{ in } L^2 \text{ as } n \rightarrow \infty \\ &\rightarrow 0 \text{ in } L^1. \end{aligned}$$

So

$$\begin{aligned}\hat{\gamma}(sl_n) &\rightarrow \int_E e^{isI(h)(x)} d\gamma(x) \\ &= \int_{\mathbb{R}} e^{ist} d(I(h)_*\gamma)(t) \\ &= \widehat{I(h)_*\gamma}(s)\end{aligned}$$

and the result follows from Bochner's Theorem. \square

Corollary 6.10. *If $f, g : [0, T] \rightarrow \mathbb{R}^n$ are in L^2 and γ is classical Wiener measure on $C_0([0, T]; \mathbb{R}^n)$ then*

$$(i) \int_{C_0} \left(\int_0^T \langle f(s), d\sigma(s) \rangle_{\mathbb{R}^n} \right) d\gamma(\sigma) = 0;$$

$$(ii) \int_{C_0} \left(\int_0^T \langle f(s), d\sigma(s) \rangle_{\mathbb{R}^n} \right)^2 d\gamma(\sigma) = \int_0^T \|f(s)\|_{\mathbb{R}^n}^2 ds = \|f\|_{L^2}^2;$$

$$(iii) \int_{C_0} \left(\int_0^T \langle f(s), d\sigma(s) \rangle_{\mathbb{R}^n} \int_0^T \langle g(s), d\sigma(s) \rangle_{\mathbb{R}^n} \right) d\gamma(\sigma) = \int_0^T \langle f(s), g(s) \rangle_{\mathbb{R}^n} ds = \langle f, g \rangle_{L^2}.$$

Proof. (i) Set $h(t) := \int_0^T f(s) ds$, so $h \in L_0^{2,1}$ and, by definition, $\int_0^T \langle f(s), d\sigma(s) \rangle = I(h)(\sigma)$. Therefore,

$$\begin{aligned}\int_{C_0} \int_0^T \langle f(s), d\sigma(s) \rangle_{\mathbb{R}^n} d\gamma(\sigma) &= \int_{C_0} I(h)(\sigma) d\gamma(\sigma) \\ &= \int_{\mathbb{R}} t d\gamma_{\langle h, - \rangle}^H(t) \text{ by Proposition 6.9} \\ &= 0 \text{ by the symmetry of } \gamma\end{aligned}$$

(ii) Recall that $\|I(h)\|_{L^2}^2 = \|h\|_{L_0^{2,1}}^2 = \|f\|_{L^2}^2$ by construction.

(iii) Follows from (ii) by the polarization identity for inner product spaces:

$$\langle a, b \rangle = \frac{\|a + b\|^2 - \|a - b\|^2}{4}.$$

We have a map $L^2([0, T]; \mathbb{R}) \rightarrow L^2(C_0, \gamma; \mathbb{R})$ given by

$$f \mapsto \left(\sigma \mapsto \int_0^T \langle f(s), d\sigma(s) \rangle_{\mathbb{R}^n} \right),$$

an isometry onto its image. \square

6.2 The Structure of Gaussian Measures

Definition 6.11. If $\{\Omega, \mathcal{F}, \mathbb{P}\}$ is a probability space, G a separable Banach space, and $f : \Omega \rightarrow G$, we say that f is a *Gaussian random variable* (or *random vector*) if

(i) f is measurable;

(ii) $f_*\mathbb{P}$ is a Gaussian measure on G .

Example 6.12. $I(h) : E \rightarrow \mathbb{R}$ is a Gaussian random variable on $\{E, \mathcal{B}(E), \gamma\}$ if $i : H \rightarrow E$ is an AWS and $h \in H$, by Proposition 6.9.

Remark 6.13. Let $\{\Omega, \mathcal{F}, \mu\}$ be a measure space, $\{X, d\}$ a metric space, and $f_j, g : \Omega \rightarrow X$ measurable functions for $j \in \mathbb{N}$. $f_j \rightarrow g$ *almost everywhere/almost surely/with probability 1* means that there is a set $Z \in \mathcal{F}$ with $\mu(Z) = 0$ such that $f_j(x) \rightarrow g(x)$ as $j \rightarrow \infty$ for all $x \notin Z$. Convergence almost everywhere is *not* implied by L^2 convergence: consider for example the sequence of functions

$$\chi_{[0,1]}, \chi_{[0,1/2]}, \chi_{[1/2,1]}, \chi_{[0,1/4]}, \chi_{[1/4,1/2]}, \chi_{[1/2,3/4]}, \dots,$$

which converges to 0 in L^2 but not almost surely. However, if the f_j are dominated then convergence almost surely implies L^2 convergence.

Lemma 6.14. Let $\{\Omega, \mathcal{F}, \mathbb{P}\}$ be a probability space and $f_j : \Omega \rightarrow \mathbb{R}$ a sequence of Gaussian random variables such that $f_j \rightarrow 0$ almost surely as $j \rightarrow \infty$. Then $f_j \rightarrow 0$ in L^2 . In particular, every Gaussian \mathbb{R} -valued random variable lies in $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$.

Proof. Set $\gamma_j := (f_j)_* \mathbb{P}$ on \mathbb{R} .

$$\begin{aligned} \widehat{\gamma}_j(s) &= \int_{\mathbb{R}} e^{ist} d\gamma_j(t) \\ &= \int_{\Omega} e^{isf_j(\omega)} d\mathbb{P}(\omega) \\ &\rightarrow 1 \text{ as } j \rightarrow \infty \text{ by DCT} \end{aligned}$$

So $\widehat{\gamma}_j(s) \rightarrow 1$ for all $s \in \mathbb{R}$ (\spadesuit). Now,

$$d\gamma_j(t) = \frac{1}{\sqrt{2\pi}} c_j^{1/2} e^{-\frac{1}{2}c_j t^2} dt$$

for some $c_j > 0$ if $f_j \not\equiv 0$, so $\widehat{\gamma}_j(s) = e^{-\frac{1}{2}c_j^{-1}s^2}$ for $s \in \mathbb{R}$, by Lemma 3.9. Therefore, $c_j^{-1} \rightarrow 0$ as $j \rightarrow \infty$ by (\spadesuit). But

$$\begin{aligned} \|f_j\|_{L^2}^2 &= \int_{\mathbb{R}} t^2 d\gamma_j(t) \\ &= \int_{\mathbb{R}} \frac{c_j^{1/2} t^2}{\sqrt{2\pi}} e^{-\frac{1}{2}c_j t^2} dt \\ &= c_j^{-1/2} \text{ by Lemma 3.1} \\ &\rightarrow 0 \end{aligned} \quad \square$$

Remark 6.15. From the proof we saw that if $f : \Omega \rightarrow \mathbb{R}$ is a Gaussian random variable then $f \in L^2$ and $\widehat{f_* \mathbb{P}}(s) = e^{-\frac{1}{2}s^2 \|f\|_{L^2}^2}$. Cf. Theorem 3.13.

Theorem 6.16. (Structure Theorem for Gaussian Measures — Kallianpur, Sato, Stefan, Dudley-Feldman-LeCamm 1977.) *Let γ be a strictly positive Gaussian measure on a separable Banach space E . Then there exists a separable Hilbert space $\{H, \langle \cdot, \cdot \rangle_H\}$ and an $i : H \rightarrow E$ such that $i : H \rightarrow E$ is an AWS with $\gamma = i_*(\gamma^H)$.*

Remark 6.17. The Structure Theorem tells us that all (centred, non-degenerate) Gaussian measures on separable Banach spaces arise as the push-forward of the canonical Gaussian CSM on some separable Hilbert space. Put another way, the AWS construction is the only way to obtain a Gaussian measure on a separable Banach space.

Proof of Theorem 6.16. We must construct H and i . Let $\ell \in E^*$, $\ell \neq 0$. Then $\ell \in L^2$ by Remark 6.15. Let $j : E^* \rightarrow L^2(E, \gamma; \mathbb{R})$ be the projection $\ell \mapsto [\ell]$. Set $H := \overline{j(E^*)}$ with $\langle \cdot, \cdot \rangle_H := \langle \cdot, \cdot \rangle_{L^2}$. Consider j as a map $E^* \rightarrow H$. By definition, this is linear with dense range. So see that it is continuous, let $\ell_n \rightarrow \ell$ in E^* as $n \rightarrow \infty$. Then

- (i) $\ell_n - \ell \rightarrow 0$ in E^* and $\ell_n - \ell$ is a Gaussian random variable on $\{E, \gamma\}$;
- (ii) $\ell_n(x) - \ell(x) \rightarrow 0$ for all $x \in E$, so $\ell_n - \ell \rightarrow 0$ (almost) surely.

Therefore, by Lemma 6.14, $\ell_n - \ell \rightarrow 0$ in L^2 , and so j is continuous. (Note: this argument shows that j is continuous from E^* with the weak-* topology to L^2 .) Now define $i := j^* : H \cong H^* \rightarrow E^{**}$ by $i(h)(\ell) := \langle h, j(\ell) \rangle_H$ for $h \in H$, $\ell \in E^*$ (\dagger).

Case 1. Suppose that E is reflexive, so the natural map $k : E \rightarrow E^{**}$ given by $k(x)(\ell) := \ell(x)$ for $x \in E$, $\ell \in E^*$, is surjective (and so is an isometry). Then we get $i : H \rightarrow E$ defined by (\dagger), which is equivalent to $\ell(i(h)) = \langle h, j(\ell) \rangle_H$ for $\ell \in E^*$, $h \in H$ (\ddagger).

Case 2. If E is not reflexive observe that the continuity of $j : (E^*, w^*) \rightarrow H$ was proved above. Use the theorem that $(E^*, w^*) \cong E$. Again, we get i satisfying (\ddagger). We now have three checks to perform:

- (i) $i_*\gamma^H = \gamma$. Observe that

$$\begin{aligned} \widehat{i_*\gamma^H}(\ell) &= \widehat{\gamma^H}(i^*(\ell)) \\ &= \widehat{\gamma^H}(j(\ell)) \\ &= e^{-\frac{1}{2}\|j(\ell)\|_H^2} \text{ by Proposition 5.6} \\ &= e^{-\frac{1}{2}\|\ell\|_{L^2}^2} \text{ by the definition of } j \end{aligned}$$

But $\widehat{\gamma}(\ell) = e^{-\frac{1}{2}\|\ell\|_{L^2}^2}$, so $i_*\gamma^H = \gamma$ by the Extended Bochner Theorem.

- (ii) i is injective. This is easy, since j has dense range.
- (iii) i has dense range. It is enough to show that j is injective. Suppose that $\ell \in \ker j$, so $j(\ell) = 0$. Then $\ell = 0$ almost surely. But if $\ell \neq 0 \in E^*$ then $\ker \ell$ is a proper closed subspace of E , so there exist $x \in E$ and $r > 0$ with $B_r(x) \cap \ker \ell = \emptyset$. So $\gamma(B_r(s)) = 0$, since $\ell = 0$ almost surely, which contradicts the strict positivity of γ . \square

Theorem 6.18. (Uniqueness of Abstract Wiener Spaces.) *Suppose that $i : H \rightarrow E$ and $i_0 : H_0 \rightarrow E$ are abstract Wiener spaces with the same measure γ on E . Then there exists a unique orthogonal $U : H_0 \rightarrow H$ ($U^*U = UU^* = \text{id}$) such that $i \circ U = i_0$:*

$$\begin{array}{ccc} H_0 & \xrightarrow{U} & H \\ & \searrow i_0 & \swarrow i \\ & & E \end{array}$$

Proof. Take $j : E^* \rightarrow H$ and $j_0 : E^* \rightarrow H_0$ as usual. We proved that $\|j(\ell)\|_H = \|\ell\|_{L^2} = \|j_0(\ell)\|_{H_0}$ for $\ell \in E^*$ in Theorem 6.3. Define $W_0 : j(E^*) \rightarrow j_0(E^*)$ by $W_0(j(\ell)) := j_0(\ell)$. This is linear and well-defined since j is injective. Also, for all $h \in j(E^*)$, $\|W_0(h)\|_{H_0} = \|h\|_H$.

Therefore, since $j(E^*)$ is dense in H , there is a unique continuous linear $W : H \rightarrow H_0$ extending W_0 . Moreover, for all $h \in H$, $\|W(h)\|_{H_0} = \|h\|_H$. Also, W is surjective since its image contains the dense subspace $j_0(E^*)$. Therefore, W is a norm-preserving isometry, so $W^*W = WW^* = \text{id}$. Also,

$$\begin{array}{ccc} H & \xrightarrow{W} & H_0 \\ & \swarrow j & \nearrow j_0 \\ & E^* & \end{array}$$

So take $U = W^* : H_0 \rightarrow H$, so since $W \circ j = j_0$, $i \circ U = i_0$ as required.

For uniqueness, note that

$$\begin{aligned} i \circ U' = i_0 &\implies (U')^* \circ j = j_0 \\ &\implies (U')^* = W \text{ since } \overline{j(E^*)} = H \\ &\implies U' = U \end{aligned}$$

□

Example 6.19. Classical Wiener Space. We used the inclusion $i : L_0^{2,1} \hookrightarrow C_0$. Other authors use $i_0 : H_0 := L^2([0, T]; \mathbb{R}) \rightarrow C_0$, where $i_0(h)(t) = \int_0^t h(s) ds$ for $0 \leq t \leq T$, $h \in H_0$. We have $U : L^2 \rightarrow L_0^{2,1}$ as $U(h)(t) = \int_0^t h(s) ds$ for $0 \leq t \leq T$, $h \in L^2$.

7 The Cameron-Martin Formula: Quasi-Invariance of Gaussian Measures

Let $i : H \rightarrow E$ be an AWS with measure γ . Consider $T_h : E \rightarrow E$ given by $T_h(x) = x + i(h)$ for $h \in H$. Suppose that $\dim E = n$ and consider γ^n on \mathbb{R}^n with $H = E = \mathbb{R}^n$ and $i = \text{id}$. Recall that

$$\gamma^n(A) := (2\pi)^{-n/2} \int_A e^{-\|x\|^2/2} dx$$

for $A \subseteq \mathbb{R}^n$ Borel. If $h \in \mathbb{R}^n$ then

$$\begin{aligned} (T_h)_*(\gamma^n)(A) &= \gamma^n(T_h^{-1}(A)) \\ &= (2\pi)^{-n/2} \int_{T_h^{-1}(A)} e^{-\|x\|^2/2} dx \\ &= (2\pi)^{-n/2} \int_A e^{-\|y-h\|^2/2} dy \\ &= \int_A e^{\langle h, y \rangle - \frac{1}{2}\|h\|^2} d\gamma^n(y). \end{aligned}$$

Therefore, $(T_h)_*(\gamma^n) = e^{\langle h, \cdot \rangle - \frac{1}{2}\|h\|^2} \gamma^n$. Thus we have

Proposition 7.1. $(T_h)_*\gamma^n \approx \gamma^n$ with Radon-Nikodym derivative

$$\frac{d(T_h)_*\gamma^n}{d\gamma^n}(x) = e^{\langle h, x \rangle_{\mathbb{R}^n} - \frac{1}{2}\|h\|_{\mathbb{R}^n}^2}.$$

Recall that if μ, ν on $\{X, \mathcal{A}\}$ are such that $\mu(A) = 0 \implies \nu(A) = 0$ then we write $\nu \prec \mu$ and say that ν is *absolutely continuous* with respect to μ . The Radon-Nikodym Theorem then says that there exists a function $\frac{d\nu}{d\mu} : X \rightarrow \mathbb{R}_{\geq 0}$ such that $\nu = \frac{d\nu}{d\mu}\mu$, i.e.

$$\nu(A) = \int_A \frac{d\nu}{d\mu}(x) d\mu(x)$$

for all $A \in \mathcal{A}$. If $\mu \prec \nu$ and $\nu \prec \mu$ then we write $\mu \approx \nu$, say μ and ν are *equivalent*, and we have $\frac{d\mu}{d\nu}(x) = (\frac{d\nu}{d\mu}(x))^{-1}$ almost everywhere.

Proposition 7.2. If μ is a probability measure (or, indeed, just finite) on a separable Banach space E , define $T_v : E \rightarrow E : x \mapsto x + v$ for a choice of $v \in E$. Then for all $\ell \in E^*$,

$$(\widehat{(T_v)_*(\mu)})(\ell) = e^{i\ell(v)} \widehat{\mu}(\ell).$$

Proof.

$$\begin{aligned} (\widehat{(T_v)_*(\mu)})(\ell) &= \int_E e^{i\ell(x)} d(T_v)_*(\mu)(x) \\ &= \int_E e^{i\ell(y+v)} d\mu(y) \\ &= e^{i\ell(v)} \widehat{\mu}(\ell). \end{aligned}$$

□

Lemma 7.3. For any AWS $i : H \rightarrow E$ with measure γ ,

(i) for $h \in H$, $e^{I(h)} \equiv e^{\langle h, - \rangle_{\tilde{H}}} \in L^p$ for all $1 \leq p < \infty$;

(ii) for all $\rho, z \in \mathbb{C}$, $g, h \in H$,

$$\int_E e^{\rho \langle g, - \rangle_{\tilde{H}}(x) + z \langle h, - \rangle_{\tilde{H}}(x)} d\gamma(x) = e^{\frac{1}{2}\rho^2 \|g\|_H^2 + \frac{1}{2}z^2 \|h\|_H^2 + \rho z \langle g, h \rangle_H}.$$

Proof. (i) Take $h \neq 0$, otherwise trivial. We know $I(h) \in L^2$ and Proposition 6.9 implies that $I(h)_* \gamma = \gamma_{\langle h, - \rangle_H}^H$. Therefore,

$$\begin{aligned} \int_E (e^{\langle h, - \rangle_{\tilde{H}}})^p d\gamma &= \int_{\mathbb{R}} t^p d(I(h)_* \gamma)(t) \\ &= \int_{\mathbb{R}} t^p d\gamma_{\langle h, - \rangle_H}^H(t) \\ &< \infty \text{ for } 1 \leq p < \infty \end{aligned}$$

since $d\gamma_{\langle h, - \rangle_H}^H(t) = \frac{1}{N} e^{-ct^2} dt$ for some $N, c > 0$.

(ii) If $\rho = ai$, $z = bi$ for some $a, b \in \mathbb{R}$, we have the desired result, since $h \mapsto I(h)$ is linear and so $\int_E e^{\rho \langle g, - \rangle_{\tilde{H}}(x) + z \langle h, - \rangle_{\tilde{H}}(x)} d\gamma(x) = \int_E e^{i \langle ag + bh, - \rangle_{\tilde{H}}} d\gamma$ and $ag + bh \in H$ since H is a real Hilbert space. So $\int_E e^{\rho \langle g, - \rangle_{\tilde{H}}(x) + z \langle h, - \rangle_{\tilde{H}}(x)} d\gamma(x) = e^{-\frac{1}{2} \|ag + bh\|_H^2}$ from before, as required.

Next fix $\rho = ai$. Both sides are analytic in $z \in \mathbb{C}$ (see below) and agree for $z \in i\mathbb{R}$, and so agree for all $z \in \mathbb{C}$. Next fix $z \in \mathbb{C}$ and observe that both sides are analytic in $\rho \in \mathbb{C}$ and agree for $\rho \in i\mathbb{R}$, and so agree for all $\rho \in \mathbb{C}$. \square

Remark 7.4. Why do we have analyticity above? Consider a measure space $\{\Omega, \mathcal{F}, \mu\}$. Let $F : \mathbb{C} \times \Omega \rightarrow \mathbb{C}$ be (jointly) measurable and $F(z, \omega)$ analytic in z for almost all $\omega \in \Omega$. When is $\int_{\Omega} F(z, \omega) d\mu(\omega)$ analytic in $z \in \mathbb{C}$?

Take a piecewise C^1 closed curve $\sigma : [0, T] \rightarrow \mathbb{C}$, $\sigma(0) = \sigma(T)$, parameterizing a closed contour \mathcal{C} . By Fubini's Theorem,

$$\int_{\mathcal{C}} \int_{\Omega} F(z, \omega) d\mu(\omega) dz = \int_{\Omega} \int_{\mathcal{C}} F(z, \omega) dz d\mu(\omega) = 0$$

by Cauchy's Theorem. This gives analyticity by Morera's Theorem. But in order to apply Fubini's Theorem we must have

$$\int_0^T \int_{\Omega} |F(\sigma(t), \omega)| |\dot{\sigma}(t)| d\mu(\omega) dt < \infty.$$

This is at most $\text{length}(\sigma) \int_{\Omega} \sup_{z \in \mathcal{C}} |F(z, \omega)| d\mu(\omega)$, where $\text{length}(\sigma) := \int_0^T |\dot{\sigma}(t)| dt$. So we are all right if $\omega \mapsto \sup_{z \in K} |F(z, \omega)|$ is in $L^1(\Omega, \mu; \mathbb{R})$ for all compact $K \subset \mathbb{C}$. But this does not hold in our case!

For us, with, say, fixed ρ ,

$$\begin{aligned} |F(z, \omega)| &= e^{\text{Re } \rho \langle g, - \rangle_{\tilde{H}}(\omega) + \text{Re } z \langle h, - \rangle_{\tilde{H}}(\omega)} \\ &= e^{\langle (\text{Re } \rho)g + (\text{Re } z)h, - \rangle_{\tilde{H}}(\omega)} \end{aligned}$$

for $\omega \in E$, and

$$\int_0^T \int_E |\dot{\sigma}(t)| |F(\sigma(t), \omega)| d\gamma(\sigma) dt = \int_0^T \int_E e^{\langle k(t), - \rangle_{\tilde{H}(\omega)}} d\gamma(\omega) |\dot{\sigma}(t)| dt < \infty,$$

where $k(t) = (\operatorname{Re} \rho)g + (\operatorname{Re} \sigma(t))h \in H$ as usual.

$$\int_E e^{\langle k(t), - \rangle_{\tilde{H}}} d\gamma = \int_{\mathbb{R}} e^s d\gamma_{\langle k(t), - \rangle_H}^H(s) < \infty.$$

Theorem 7.5. (Cameron-Martin Formula.) *For an AWS $i : H \rightarrow E$ with measure γ , let $T_h : E \rightarrow E$ be $T_h(x) := x + i(h)$ for $h \in H$. Then $(T_h)_*\gamma \approx \gamma$ with*

$$(T_h)_*\gamma = e^{\langle h, - \rangle_{\tilde{H}} - \frac{1}{2}\|h\|_H^2} \gamma.$$

Remark 7.6. The Cameron-Martin Theorem is the analogue of Proposition 7.1 for translations by elements of the dense subspace $i(H) \subseteq E$.

Proof of Theorem 7.5. Set $\gamma_h := (T_h)_*\gamma$. By Proposition 7.2, for $\ell \in E^*$,

$$\begin{aligned} \hat{\gamma}_h &= e^{\sqrt{-1}\ell(i(h))} \hat{\gamma}(\ell) \\ &= e^{\sqrt{-1}\ell(i(h))} e^{-\frac{1}{2}\|j(\ell)\|_H^2} \\ &= e^{\sqrt{-1}\langle j(\ell), h \rangle_{\tilde{H}} - \frac{1}{2}\|j(\ell)\|_H^2}. \end{aligned}$$

Now set $\tilde{\gamma} := e^{\langle h, - \rangle_{\tilde{H}} - \frac{1}{2}\|h\|_H^2} \gamma$.

$$\begin{aligned} \hat{\gamma}_h(\ell) &= \int_E e^{\sqrt{-1}\ell(x)} d\tilde{\gamma}(x) \\ &= \int_E e^{\sqrt{-1}\ell(x)} e^{\langle h, - \rangle_{\tilde{H}} - \frac{1}{2}\|h\|_H^2} d\gamma(x) \\ &= e^{-\frac{1}{2}\|h\|_H^2} \int_E e^{\sqrt{-1}\langle j(\ell), h \rangle_{\tilde{H}} + \langle h, - \rangle_{\tilde{H}}(x)} d\gamma(x) \\ &= e^{-\frac{1}{2}\|h\|_H^2} e^{-\frac{1}{2}\|j(\ell)\|_H^2 + \frac{1}{2}\|h\|_H^2 + \sqrt{-1}\langle j(\ell), h \rangle_H} \\ &= \hat{\gamma}_h(\ell) \end{aligned}$$

Therefore, Bochner's Theorem for infinite dimensions implies that $\gamma_h = \tilde{\gamma}$. □

Theorem 7.7. (Integrated Cameron-Martin.) *If $F : E \rightarrow \mathbb{R}$ (or $E \rightarrow$ any separable Banach space) is measurable and $h \in H$ then*

$$\int_E F(x + i(h)) d\gamma(x) = \int_E F(x) e^{\langle h, - \rangle_{\tilde{H}}(x) - \frac{1}{2}\|h\|_H^2} d\gamma(x)$$

in the sense that if one side exists, both exist and are equal.

Proof. By Theorem 7.5 and Proposition 2.4,

$$\gamma \longrightarrow (T_h)_*\gamma$$

$$\begin{array}{ccc} E & \xrightarrow{T_h} & E \\ & \searrow F \circ T_h & \downarrow F \\ & & \mathbb{R} \end{array}$$

□

Remarks 7.8. (i) Consider $t \mapsto th : \mathbb{R} \rightarrow H$. We get

$$\int_E F(x + ti(h)) d\gamma(x) = \int_E F(x) e^{t\langle h, - \rangle \sim(x) - \frac{1}{2}t^2 \|h\|_H^2} d\gamma(x).$$

Formally differentiate at $t = 0$:

$$\int_E DF(x)(i(h)) d\gamma(x) = \int_E F(x) \langle h, - \rangle \sim(x) d\gamma(x).$$

If F is a “nice” differentiable function with derivative $DF : E \rightarrow \mathbb{L}(E; \mathbb{R})$, we have the above integration by parts formula.

(ii) In \mathbb{R} ,

$$\int_{\mathbb{R}} f'(x)v(x) dx = - \int_{\mathbb{R}} f(x)v'(x) dx$$

if f and v are “nice” (“vanishing at ∞ ”), $f, v, f', v' \in L^2(\mathbb{R}; \mathbb{R})$. In \mathbb{R}^n , for $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and a vector field $V : \mathbb{R}^n \rightarrow \mathbb{R}^n$,

$$\begin{aligned} \int_{\mathbb{R}^n} Df(x)(V(x)) dx &= \int_{\mathbb{R}^n} \langle \nabla f(x), V(x) \rangle_{\mathbb{R}^n} dx \\ &= - \int_{\mathbb{R}^n} \operatorname{div} V(x) f(x) dx \end{aligned}$$

For us the vector field is $V : E \rightarrow E : x \mapsto i(h)$ for all $x \in E$.

$$\text{“div } V(x)\text{”} = -\langle h, - \rangle \sim(x)$$

We want to allow more general vector fields $V : E \rightarrow E$. In the classical case these will be “stochastic processes”.

Remark 7.9. The Cameron-Martin Theorem says that γ is quasi-invariant under translations by elements in the image of i . The converse is also true: γ is quasi-invariant under $x \mapsto x + v \iff v \in i(H)$.

Theorem 7.10. *If H is an infinite-dimensional separable Hilbert space then the canonical Gaussian CSM on H , $\{\gamma^H\}$, is not a measure on H .*

Proof. Suppose not, so that $\gamma^H = \gamma$ is a measure on H . Then $i = \operatorname{id} : H \rightarrow H$ is γ -radonifying, and so an AWS. So, by the Cameron-Martin Theorem, γ is quasi-invariant under all $h \in H$. Thus $\dim H < \infty$ by Theorem 1.15. \square

Remarks 7.11. For $i : H \rightarrow E$ an AWS,

- (i) It is possible to show that $i(H)$ is Borel measurable in E and has measure 0.
- (ii) L. Gross proved that $\exists i_0 : H \rightarrow E_0$, also an AWS, and $k \in \mathbb{L}(E_0; E)$ injective and compact such that $k \circ i_0 = i$:

$$\begin{array}{ccc} H & \xrightarrow{i} & E \\ & \searrow i_0 & \uparrow k \\ & & E_0 \end{array}$$

So γ “lives” on $E_0 \subseteq E$. (k compact means that $\overline{k(\text{bounded set})}$ is compact.)

Example 7.12. Classical case, $E = C_0$, $E_0 =$ closure of $L_0^{2,1}$ in the norm

$$\|\sigma\|_{0+\alpha} := \sup_{s,t \in [0,T], s \neq t} \frac{|\sigma(s) - \sigma(t)|}{|s - t|^\alpha}$$

for any $0 < \alpha < \frac{1}{2}$.

8 Stochastic Processes and Brownian Motion in \mathbb{R}^n

8.1 Stochastic Processes

Definition 8.1. A *stochastic process* indexed by a set S with state space a measurable space $\{X, \mathcal{A}\}$ is a map $z : S \times \Omega \rightarrow X$ for some probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$ such that for all $s \in S$, the map $\Omega \rightarrow X : \omega \mapsto z_s(\omega) := z(s, \omega)$ is measurable.

Often $S = [0, T]$ for some $T > 0$, or $S = [0, \infty)$. A stochastic process is then a family of maps/paths $S \rightarrow X : s \mapsto z_s(\omega)$ parametrized by $\omega \in \Omega$.

In the Kolmogorov model of probability theory, the probability that our system or process behaves in a certain way is $\mathbb{P}\{\omega \in \Omega | s \mapsto z_s(\omega) \text{ behaves that way}\}$ for a suitably chosen stochastic process.

Example 8.2. If $A_i \in \mathcal{A}$ for $i = 1, \dots, k$ and $s_1, \dots, s_k \in S$ then the probability that the process has value in A_i at $s = s_i$ for each i is $\mathbb{P}\{\omega \in \Omega | z_{s_i}(\omega) \in A_i \text{ for } i = 1, \dots, k\}$, which we often write as $\mathbb{P}\{z_{s_i} \in A_i \text{ for } i = 1, \dots, k\}$ or $\mathbb{P}_{\mathbf{s}}(A_1 \times \dots \times A_k)$, where $\mathbb{P}_{\mathbf{s}}$ is the push-forward measure $(z_{\mathbf{s}})_*(\mathbb{P})$ on X^k , where

$$\begin{aligned} z_{\mathbf{s}} : \Omega &\rightarrow X^k \\ \omega &\mapsto (z_{s_1}(\omega), \dots, z_{s_k}(\omega)), \end{aligned}$$

and $\mathbf{s} := (s_1, \dots, s_k)$. These $\{\mathbb{P}_{\mathbf{s}} | \mathbf{s} \subseteq S \text{ finite}\}$ are probability measures on X^k , called the *finite-dimensional distributions* of the process.

Example 8.3. If $i : H \rightarrow E$ is an AWS with measure γ , consider $z : H \times E \rightarrow \mathbb{R}$, $\{E, \mathcal{B}(E), \gamma\}$ our measure space, given by $z(h, \omega) = \langle h, - \rangle_{\tilde{H}}(\omega)$ so $S = H$ here. (Strictly, we need to choose a representative of the class of $\langle h, - \rangle^{\sim}$ in L^2 .)

Definition 8.4. If S, X are topological spaces (and $\mathcal{A} = \mathcal{B}(X)$) we say that a stochastic process z is *continuous* (or *sample continuous*) if the map $S \rightarrow X : s \mapsto z_s(\omega)$ is continuous for all $\omega \in \Omega$. (Some authors allow almost all $\omega \in \Omega$.)

Example 8.5. Let $\Omega = C_0 = C_0([0, T]; \mathbb{R}^n) = \{\text{continuous paths in } \mathbb{R}^n \text{ starting at } 0\}$, $\mathbb{P} = \text{Wiener measure}$, $X = \mathbb{R}^n$, $\mathcal{A} = \mathcal{B}(\mathbb{R}^n)$, $S = [0, T]$. Any measurable vector field $V : C_0 \rightarrow C_0$ determines a sample continuous process $z : [0, T] \times C_0 \rightarrow \mathbb{R}^n$ by $z_s(\sigma) = V(\sigma)(s)$ for $\sigma \in C_0$ and $s \in [0, T]$.

Simplest example: $V = \text{id}$, so $V(\sigma) = \sigma$ for all $\sigma \in C_0$. Then $z_s(\sigma) = \sigma(s)$ for $0 \leq s \leq T$, which we call the *canonical process* on $C_0([0, T]; \mathbb{R}^n)$.

Exercise 8.6. Conversely, if $z : [0, T] \times \Omega \rightarrow X$ is a continuous process, with X a separable Banach space, we get $\Phi : \Omega \rightarrow C([0, T]; X)$ given by $\Phi(\omega)(t) = z_t(\omega) \in X$. Check that this is measurable. Hint: Use Lemma 5.18 and try the special case $X = \mathbb{R}$.

Definition 8.7. The *law* \mathcal{L}_z of z is the push-forward measure $\mathcal{L}_z := \Phi_*(\mathbb{P})$ on $C([0, T]; X)$. (If $z_0(\omega) = 0$ for all $\omega \in \Omega$ we can use $C_0([0, T]; X)$.)

Remark 8.8. The canonical process $[0, T] \times C([0, T]; X) \rightarrow X$ using the measure \mathcal{L}_z on $C([0, T]; X)$ has the same finite-dimensional distributions as $z : [0, T] \times \Omega \rightarrow X$. Consequently, the finite-dimensional distributions of z determine its law.

Definition 8.9. A *Brownian motion* (or BM) on \mathbb{R}^n is a stochastic process $B : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ such that

- (i) $B_0(\omega) = 0$ for all $\omega \in \Omega$;
- (ii) B is sample continuous;
- (iii) \mathcal{L}_B is Wiener measure on $C_0([0, T]; \mathbb{R}^n)$.

Example 8.10. Canonical BM with $\Omega = C_0$, $\mathbb{P} =$ Wiener measure.

Remark 8.11. We could write $[0, \infty)$ instead of $[0, T]$ but then we would have to take care with condition (iii): it is enough to say that $B : [0, \infty) \times \Omega \rightarrow \mathbb{R}^n$ is a BM on \mathbb{R}^n if for all $T > 0$, $B|_{[0, T] \times \Omega}$ is a BM on \mathbb{R}^n .

Definition 8.12. If $h \in H := L^{2,1}([0, T]; \mathbb{R}) \subset C_0$ and B is a BM on \mathbb{R}^n , write $\int_0^T \langle \dot{h}(s), dB_s \rangle_{\mathbb{R}^n} : \Omega \rightarrow \mathbb{R}$ for the composition

$$\begin{aligned} \Omega &\xrightarrow{\Phi} C_0 \xrightarrow{I(h)=\langle h, - \rangle^\sim} \mathbb{R} \\ \omega &\mapsto B_s(\omega) \end{aligned}$$

which makes sense since Φ preserves sets of measure 0. Then if $f \in L^2([0, T]; \mathbb{R})$,

$$\int_0^T \langle f(s), dB_s \rangle := I \left(\int_0^\cdot f(s) ds \right) \circ \Phi : \Omega \rightarrow \mathbb{R}.$$

Definition 8.13. For a function f on a probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$, write $\mathbb{E}f := \int_{\Omega} f(x) d\mathbb{P}(x)$ for the *expectation* of f .

Theorem 8.14. For B a BM on \mathbb{R}^n and $f \in L^2([0, T]; \mathbb{R})$, $\int_0^T \langle f(s), dB_s \rangle : \Omega \rightarrow \mathbb{R}$ is in $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ and

$$(i) \quad e^{i \int_0^T \langle f(s), dB_s \rangle} = e^{-\frac{1}{2} \int_0^T |f(s)|^2 ds} = e^{-\frac{1}{2} \|f\|_{L^2}^2};$$

$$(ii) \quad \text{for } f, g \in L^2([0, T]; \mathbb{R}), \quad \mathbb{E} \int_0^T \langle f(s), dB_s \rangle \int_0^T \langle g(s), dB_s \rangle = \int_0^T \langle f(s), g(s) \rangle ds = \langle f, g \rangle_{L^2}$$

$$(iii) \quad \mathbb{E} \int_0^T \langle f(s), dB_s \rangle = 0.$$

Proof. This follows from Corollary 6.10 and the push-forward theorem. □

8.2 Construction of Itô's Integral

We want to construct the Itô integral

$$\int_0^T \langle a_s, dB_s \rangle : \Omega \rightarrow \mathbb{R},$$

where $a : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ is a process. In the canonical version,

$$\sigma \mapsto \int_0^T \langle a_s(\sigma), d\sigma_s \rangle.$$

Note that in the Paley-Wiener integral $\int_0^t \langle f(s), dB_s \rangle$ we have f not dependent upon $\omega \in \Omega$. These are “constant vector fields”. We want to give a more concrete definition that includes time evolution.

Definition 8.15. Let $\{\Omega, \mathcal{F}, \mathbb{P}\}$ be a probability space. A *filtration* is a family of σ -algebras on Ω , $\{\mathcal{F}_t | t \in [0, T]\}$, such that

- for all $t \in [0, T]$, $\mathcal{F}_t \subseteq \mathcal{F}$; and
- $0 \leq s \leq t \leq T \implies \mathcal{F}_s \subseteq \mathcal{F}_t$.

Example 8.16. Given a process $z : [0, T] \times \Omega \rightarrow X$, with $\{X, \mathcal{A}\}$ any measurable space, define $\mathcal{F}_t := \mathcal{F}_t^z = \sigma\{z_r : \Omega \rightarrow X | 0 \leq r \leq t\}$, the “events up to time t ” or “the past at time t ”. \mathcal{F}_*^z is called the *natural filtration* of z .

Example 8.17. If $\Omega = C_0([0, T]; \mathbb{R}^n)$ and z is canonical ($z_s(\omega) = \omega(s)$), then if $0 \leq s_1 \leq s_2 \leq \dots \leq s_k \leq t$ and $A_1, \dots, A_k \in \mathcal{B}(\mathbb{R}^n)$,

$$\{\sigma \in C_0 | \sigma(s_j) \in A_j, 1 \leq j \leq k\} \in \mathcal{F}_t^z.$$

To define Itô’s integral $\int_0^T \langle a_s, dB_s \rangle$ for $a : [0, T] \times \Omega \rightarrow \mathbb{R}^n$, we will need a to be “non-anticipating” or “adapted” to some filtration of \mathcal{F} , usually \mathcal{F}_*^B .

Definition 8.18. A process $a : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ is *adapted* to a filtration $\{\mathcal{F}_t | 0 \leq t \leq T\}$ if $a_t : \Omega \rightarrow \mathbb{R}^n$ is \mathcal{F}_t -measurable for all $0 \leq t \leq T$. (In general, we can replace \mathbb{R}^n by any measurable space $\{X, \mathcal{A}\}$.)

Take $n = 1$ for ease of notation.

Definitions 8.19. Given a filtration $\{\mathcal{F}_t | 0 \leq t \leq T\}$ on $\{\Omega, \mathcal{F}, \mathbb{P}\}$, a process $a : [0, T] \times \Omega \rightarrow \mathbb{R}$ is *elementary* if for all $\omega \in \Omega$, $0 \leq t \leq T$,

$$a_t(\omega) = \alpha_{-1}(\omega)\chi_{\{0\}}(t) + \sum_{j=0}^{k-1} \alpha_j(\omega)\chi_{(t_j, t_{j+1}]}(t)$$

for some partition $0 \leq t_0 < t_1 < \dots < t_k \leq T$ of $[0, T]$. (Some authors, such as [Ø], use $[t_j, t_{j+1}]$.) Here each $\alpha_j : \Omega \rightarrow \mathbb{R}$ is \mathcal{F}_{t_j} -measurable for each $j = 0, \dots, k-1$ and α_{-1} is \mathcal{F}_0 -measurable. Write $\mathcal{E}([0, T]; \mathbb{R})$ for the collection of all elementary processes $[0, T] \times \Omega \rightarrow \mathbb{R}$. By comparison with the Fundamental Theorem of Calculus, it is reasonable to define, for elementary processes $a \in \mathcal{E}([0, T]; \mathbb{R})$,

$$\int_0^T \langle a_s, dB_s \rangle(\omega) := \sum_{j=0}^{k-1} \alpha_j(\omega) (B_{t_{j+1}}(\omega) - B_{t_j}(\omega)).$$

Now we approximate more general processes by elementary processes to get integrals converging in the function space $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$.

Let B be a BM on \mathbb{R} , $B : [0, T] \times \Omega \rightarrow \mathbb{R}$. Given $0 \leq t_0 < t_1 < \dots < t_k \leq T$, set

$$\begin{aligned} \Delta_j B(\omega) &:= B_{t_{j+1}}(\omega) - B_{t_j}(\omega) \\ \Delta_j t &= t_{j+1} - t_j. \end{aligned}$$

Then

$$\begin{aligned}
\left\| \sum_{j=0}^k \alpha_j \Delta_j B \right\|_{L^2}^2 &= \int_{\Omega} \left| \sum_{j=0}^k \alpha_j(\omega) \Delta_j B(\omega) \right|^2 d\mathbb{P}(\omega) \\
&= 2 \int_{\Omega} \sum_{i < j} \alpha_j(\omega) \Delta_j B(\omega) \alpha_i(\omega) \Delta_i B(\omega) d\mathbb{P}(\omega) \\
&\quad + \int_{\Omega} \sum_i (\alpha_i(\omega) \Delta_i B(\omega))^2 d\mathbb{P}(\omega)
\end{aligned}$$

We will show that for suitable filtrations \mathcal{F}_* , e.g. $\mathcal{F}_t := \mathcal{F}_t^B$,

Proposition 8.20. *If B is a BM on \mathbb{R} and α_j is \mathcal{F}_{t_j} -measurable for each j and bounded then*

- (i) *if $i < j$, $\int_{\Omega} \alpha_i \alpha_j \Delta_i B \Delta_j B d\mathbb{P} = 0$;*
- (ii) *$\int_{\Omega} \alpha_i^2 (\Delta_i B)^2 d\mathbb{P} = (\mathbb{E} \alpha_i^2) \Delta_i t$;*
- (iii) *as an immediate consequence of (i) and (ii),*

$$\begin{aligned}
\left\| \sum_j \alpha_j \Delta_j B \right\|_{L^2}^2 &= \left\| \int_0^T a_s dB_s \right\|_{L^2}^2 \\
&= \sum_j \|\alpha_j\|_{L^2}^2 \Delta_j t \\
&= \int_0^T \|a_s\|_{L^2}^2 ds \\
&= \|a_s\|_{L^2([0,T] \times \Omega; \mathbb{R})}^2.
\end{aligned}$$

Assuming Proposition 8.20,

Theorem 8.21. *For an elementary bounded process $a : [0, T] \times \Omega \rightarrow \mathbb{R}$,*

$$\left\| \int_0^T a_s dB_s \right\|_{L^2} = \|a\|_{L^2([0,T] \times \Omega; \mathbb{R})}.$$

For B and \mathcal{F}_* as above, let $\mathcal{E} := \mathcal{E}([0, T]; \mathbb{R})$ be the space of elementary bounded processes $a : [0, T] \times \Omega \rightarrow \mathbb{R}$, with norm

$$\|a\| := \sqrt{\int_0^T \mathbb{E} |a_s|^2 ds} = \|a\|_{L^2([0,T] \times \Omega; \mathbb{R})}.$$

Let $\bar{\mathcal{E}}$ be the closure of \mathcal{E} in $L^2([0, T] \times \Omega; \mathbb{R})$. Define $\mathcal{I} : \mathcal{E} \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ by $\mathcal{I}(a) := \int_0^T a_s dB_s$.

Corollary 8.22. *\mathcal{I} extends uniquely to a continuous linear $\bar{\mathcal{I}} : \bar{\mathcal{E}} \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$. This is an isometry into $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$. Write it as $\bar{\mathcal{I}}(a) = \int_0^T a_s dB_s$. So, for $a \in \bar{\mathcal{E}}$, we have the Itô isometry*

$$\left\| \int_0^T a_s dB_s \right\|_{L^2(\Omega; \mathbb{R})} = \|a\|_{L^2(\Omega \times [0,T]; \mathbb{R})},$$

i.e.,

$$\int_{\Omega} \left(\int_0^T a_s dB_s(\omega) \right)^2 d\mathbb{P}(\omega) = \int_0^T \mathbb{E}|a_s|^2 ds.$$

Proof. Lemma 6.2. □

We still need to prove Proposition 8.20 parts (i) and (ii), identify $\bar{\mathcal{E}}$, and relate the Itô and Paley-Wiener integrals.

Theorem 8.23. *Let $\{\Omega, \mathcal{F}, \mathbb{P}\}$ be a probability space and $\mathcal{A} \subset \mathcal{F}$ a σ -algebra. For $f \in L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ there exists a unique $\bar{f} \in L^1(\Omega, \mathcal{A}, \mathbb{P}|_{\mathcal{A}}; \mathbb{R})$ such that, for all $A \in \mathcal{A}$,*

$$\int_A f d\mathbb{P} = \int_A \bar{f} d\mathbb{P}.$$

If $f \in L^2$ then $\bar{f} \in L^2$ and $\bar{f} = P^{\mathcal{A}}(f)$ for $P^{\mathcal{A}} : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}) \rightarrow L^2(\Omega, \mathcal{A}, \mathbb{P}|_{\mathcal{A}}; \mathbb{R})$ the orthogonal projection. Write \bar{f} as $\mathbb{E}\{f|\mathcal{A}\}$, the conditional expectation of f given / with respect to \mathcal{A} . Then also

(i) $f \geq 0$ almost everywhere $\implies \bar{f} \geq 0$ almost everywhere;

(ii) $|\mathbb{E}\{f|\mathcal{A}\}(\omega)| \leq \mathbb{E}\{|f|\mathcal{A}\}(\omega)$ almost everywhere;

(iii) $\mathbb{E}\{-|\mathcal{A}\} : L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}) \rightarrow L^1(\Omega, \mathcal{A}, \mathbb{P}|_{\mathcal{A}}; \mathbb{R})$ is a continuous linear map with norm 1.

Proof. (Uniqueness.) Suppose that \bar{f} and \tilde{f} are \mathcal{A} -measurable and satisfy

$$\int_A f d\mathbb{P} = \int_A \bar{f} d\mathbb{P} = \int_A \tilde{f} d\mathbb{P}.$$

Set $g := \bar{f} - \tilde{f}$, which is \mathcal{A} -measurable, in L^1 , and has $\int_A g d\mathbb{P} = 0$ for all $A \in \mathcal{A}$. Thus, $g = 0$ almost surely, and so $\bar{f} = \tilde{f}$ almost surely.

(L^2 part.) $P^{\mathcal{A}} f$ satisfies our criteria for \bar{f} since

- it is \mathcal{A} -measurable;
- it is in L^2 , and so is in L^1 ;
- if $A \in \mathcal{A}$,

$$\begin{aligned} \int_A P^{\mathcal{A}} f d\mathbb{P} &= \int_{\Omega} \chi_A P^{\mathcal{A}} f d\mathbb{P} \\ &= \langle \chi_A, P^{\mathcal{A}} f \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})} \\ &= \langle P^{\mathcal{A}} \chi_A, f \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})} \text{ since } (P^{\mathcal{A}})^* = P^{\mathcal{A}} \\ &= \langle \chi_A, f \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})} \\ &= \int_A f d\mathbb{P}. \end{aligned}$$

Thus $\bar{f} = P^{\mathcal{A}} f$ by uniqueness.

(Existence.) For $f \in L^1$ and $f \geq 0$ almost everywhere, define μ_f on $\{\Omega, \mathcal{A}\}$ by $\mu_f(A) = \int_A f \, d\mathbb{P}$ for $A \in \mathcal{A}$, a measure with $\mu_f \prec \mathbb{P}|_{\mathcal{A}}$. Set $\bar{f} := \frac{d\mu_f}{d(\mathbb{P}|_{\mathcal{A}})} : \Omega \rightarrow \mathbb{R}_{\geq 0}$. This satisfies the requirements for \bar{f} since, if $A \in \mathcal{A}$,

$$\int_A \frac{d\mu_f}{d(\mathbb{P}|_{\mathcal{A}})} \, d\mathbb{P} = \int_A \frac{d\mu_f}{d(\mathbb{P}|_{\mathcal{A}})} \, d(\mathbb{P}|_{\mathcal{A}}) = \int_A d\mu_f = \mu_f(A) = \int_A f \, d\mathbb{P}$$

From this $\int_{\Omega} \bar{f} \, d\mathbb{P} = \int_{\Omega} f \, d\mathbb{P} < \infty$ so $\bar{f} \in L^1$ since $\bar{f} \geq 0$. This also gives (i). For general $f \in L^1$, write $f = f^+ - f^-$ in the usual way and take $\bar{f} = \bar{f}^+ - \bar{f}^-$. It is easy to see that this satisfies all the requirements, but we must check (ii) and (iii).

(ii) $|f(\omega)| = |f|(\omega) = f^+(\omega) + f^-(\omega)$, so $\mathbb{E}\{|f|\mid\mathcal{A}\}(\omega) = \bar{f}^+(\omega) + \bar{f}^-(\omega)$. Also, $|\mathbb{E}\{f\mid\mathcal{A}\}(\omega)| = |\bar{f}(\omega)| = |\bar{f}^+(\omega) - \bar{f}^-(\omega)|$, giving (ii), since $\bar{f}^+ \geq 0$, $\bar{f}^- \geq 0$ almost everywhere.

(iii) If $f \in L^1$,

$$\begin{aligned} \|\mathbb{E}\{f\mid\mathcal{A}\}\|_{L^1} &= \int_{\Omega} |\mathbb{E}\{f\mid\mathcal{A}\}| \, d\mathbb{P} \\ &\leq \int_{\Omega} \mathbb{E}\{|f|\mid\mathcal{A}\} \, d\mathbb{P} \text{ by (i)} \\ &= \int_{\Omega} |f| \, d\mathbb{P} \\ &= \|f\|_{L^1} \end{aligned}$$

So $\mathbb{E}\{-|\mathcal{A}\}$ is bounded linear with norm ≤ 1 . But if $f \equiv 1$, $\mathbb{E}\{f\mid\mathcal{A}\} \equiv 1$, so the norm is 1. \square

Definition 8.24. If $\theta : \Omega \rightarrow X$ is measurable, $\{X, \mathcal{A}\}$ a measure space, define $\mathbb{E}\{-|\theta\} := \mathbb{E}\{-|\sigma(\theta)\}$, where $\sigma(\theta) := \{\theta^{-1}(A) \mid A \in \mathcal{A}\}$.

Lemma 8.25. *Given a probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$, a measurable space $\{X, \mathcal{A}\}$, and $\theta : \Omega \rightarrow X$ and $f : \Omega \rightarrow \mathbb{R}$ measurable, there exists a measurable $g : X \rightarrow \mathbb{R}$ such that $\mathbb{E}\{f\mid\theta\} = g \circ \theta$ almost everywhere, and this map is unique $\theta_*(\mathbb{P})$ -almost surely.*

$$\begin{array}{ccc} \Omega & \xrightarrow{\theta} & X \\ & \searrow \mathbb{E}\{f\mid\theta\} & \swarrow g \\ & & \mathbb{R} \end{array}$$

Proof. Suppose $f \geq 0$. We have \mathbb{P}_f on $\{\Omega, \mathcal{F}\}$ given by $\mathbb{P}_f(A) := \int_A f \, d\mathbb{P}$ for $A \in \mathcal{F}$. We get a probability measure $\theta_*(\mathbb{P}_f)$ on $\{X, \mathcal{A}\}$; note that $\theta_*(\mathbb{P}_f) \prec \theta_*(\mathbb{P})$. Set $g := \frac{\theta_*(\mathbb{P}_f)}{\theta_*(\mathbb{P})} : X \rightarrow \mathbb{R}_{\geq 0}$, which is \mathcal{A} -measurable. We claim that $g \circ \theta = \mathbb{E}\{f\mid\theta\}$. To see this note that

- $g \circ \theta$ is $\sigma(\theta)$ -measurable;

- if $A \in \mathcal{A}$ then

$$\begin{aligned}
\int_{\theta^{-1}(A)} g \circ \theta \, d\mathbb{P} &= \int_A g \, d\theta_*(\mathbb{P}) \\
&= \int_A \frac{\theta_*(\mathbb{P}_f)}{\theta_*(\mathbb{P})} \, d\theta_*(\mathbb{P}) \\
&= \int_A d\theta_*(\mathbb{P}_f) \\
&= \int_{\theta^{-1}(A)} d\mathbb{P}_f \\
&= \int_{\theta^{-1}(A)} f \, d\mathbb{P}
\end{aligned}$$

- taking $A = X$ above gives $g \circ \theta \in L^1$, therefore $g \circ \theta = \mathbb{E}\{f|\theta\}$.

For general f write $f = f^+ - f^-$ as before to get $g = g^+ - g^-$. □

Remark 8.26. We write $\mathbb{E}\{f|\theta = x\}$ for $g(x) =$ “the conditional expectation of f given $\theta = x$ ” or “given $\theta(\omega) = x$ ” for $x \in X$. This gives us an intuitive way to calculate $\mathbb{E}\{f|\theta\}$ for f and θ as in Lemma 8.25:

- let x be some value of θ ;
- calculate the “average value” of f on the preimage $\theta^{-1}(x)$;
- the result of this calculation is $\mathbb{E}\{f|\theta = x\}$; call it $g(x)$;
- the conditional expectation $\mathbb{E}\{f|\theta\} : \Omega \rightarrow \mathbb{R}$ is given by

$$\mathbb{E}\{f|\theta\}(\omega) = g(\theta(\omega)) \text{ for } \omega \in \Omega.$$

Remark 8.27. If f is $\sigma(\theta)$ -measurable then $f = \mathbb{E}\{f|\theta\}$ almost surely, so there is a g with $f = g \circ \theta$ almost surely.

Example 8.28. (Weather forecasting.) Let Ω be the set of all time evolutions of all possible weather patterns. Let θ_n be the value at the n th morning, i.e.

$$\theta_n : \Omega \rightarrow X = \{\{\text{wind speeds}\} \times \{\text{rain volume}\} \times \dots\}.$$

Consider just wind speed, for instance. Let $f_n : \Omega \rightarrow \mathbb{R}$ be the windspeed at mid-day on the n th day. Given some observation in the morning, we want to forecast f_n . We need $g_n : X \rightarrow \mathbb{R}$, which tells us that if on the n th morning $x \in X$ holds then $g_n(x)$ is the windspeed at mid-day. That is, we need g_n such that $g_n \circ \theta_n$ is the “best” estimate that we can make of f_n . “Best” usually means best in the mean square, the $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ norm. Now, by Lemma 8.25 functions of the form $g_n \circ \theta_n$ in L^2 are exactly elements of $L^2(\Omega, \sigma(\theta_n), \mathbb{P}|_{\sigma(\theta_n)}; \mathbb{R})$. To get the closest of these to f_n , take $P(f_n)$, the orthogonal projection of f_n , i.e., we take $\mathbb{E}\{f_n|\sigma(\theta_n)\}$.

Remark 8.29. $\mathbb{E}\{f|\theta = x\} = \int_{\theta^{-1}(x)} f(y) \, d\mathbb{P}_x(y)$, where \mathbb{P}_x is a measure on the fibre $\theta^{-1}(x)$, known as the *disintegration of \mathbb{P}* .

Definition 8.30. Let $\{\Omega, \mathcal{F}, \mathbb{P}\}$ be a probability space.

- (i) If $\mathcal{A}, \mathcal{B} \subseteq \mathcal{F}$, we say that \mathcal{A} and \mathcal{B} are *independent*, and write $\mathcal{A} \amalg \mathcal{B}$, if $A \in \mathcal{A}, B \in \mathcal{B} \implies \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
- (ii) If for $j = 1, 2$, $\{X_j, \mathcal{A}_j\}$ are measurable spaces with $f_j : \Omega \rightarrow X_j$ measurable functions, f_1 and f_2 are *independent*, $f_1 \amalg f_2$, if $\sigma(f_1) \amalg \sigma(f_2)$.

Theorem 8.31. $f_1 \amalg f_2$ if, and only if, the product function $f_1 \times f_2 : \Omega \rightarrow X_1 \times X_2 : \omega \mapsto (f_1(\omega), f_2(\omega))$ satisfies $(f_1 \times f_2)_*(\mathbb{P}) = (f_1)_*(\mathbb{P}) \otimes (f_2)_*(\mathbb{P})$, the product of the two push-forward measures.

Recall that a measure μ on $X_1 \times X_2$ is a product $\mu_1 \otimes \mu_2$ if, and only if, for all $A_1 \in \mathcal{A}_1$ and $A_2 \in \mathcal{A}_2$, we have $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$, since μ is determined by its values on rectangles.

Proof. Assume that $f_1 \amalg f_2$ and $A_j \in \mathcal{A}_j$ for $j = 1, 2$. Then

$$\begin{aligned} (f_1 \times f_2)_*(\mathbb{P})(A_1 \times A_2) &= \mathbb{P}\{\omega \in \Omega | (f_1(\omega), f_2(\omega)) \in A_1 \times A_2\} \\ &= \mathbb{P}\{\omega \in f_1^{-1}(A_1) \cap f_2^{-1}(A_2)\} \\ &= \mathbb{P}(f_1^{-1}(A_1))\mathbb{P}(f_2^{-1}(A_2)) \text{ since } f_1 \amalg f_2 \\ &= (f_1)_*(\mathbb{P})(A_1)(f_2)_*(\mathbb{P})(A_2) \end{aligned}$$

So we have a product measure. Conversely, suppose it is the product measure. Take typical elements $f_1^{-1}(A_1) \in \sigma(f_1)$ and $f_2^{-1}(A_2) \in \sigma(f_2)$. Then

$$\begin{aligned} \mathbb{P}(f_1^{-1}(A_1) \cap f_2^{-1}(A_2)) &= (f_1 \times f_2)_*(\mathbb{P})(A_1 \times A_2) \\ &= (f_1)_*(\mathbb{P})(A_1)(f_2)_*(\mathbb{P})(A_2) \text{ by hypothesis} \\ &= \mathbb{P}(f_1^{-1}(A_1))\mathbb{P}(f_2^{-1}(A_2)), \end{aligned}$$

so $f_1 \amalg f_2$. □

Corollary 8.32. For f_1, f_2 as above with $f_1 \amalg f_2$, if $F_j : X_j \rightarrow \mathbb{R}$ are measurable for $j = 1, 2$, then

$$\mathbb{E}\{F_1(f_1(-))F_2(f_2(-))\} = \mathbb{E}F_1(f_1(-))\mathbb{E}F_2(f_2(-))$$

provided $F_1 \circ f_1, F_2 \circ f_2 \in L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$.

Proof. The left-hand side is

$$\int_{X_1 \times X_2} F_1(x)F_2(y) d(f_1 \times f_2)_*(\mathbb{P})(x, y) = \int_{X_1} F_1(x) d(f_1)_*(\mathbb{P})(x) \int_{X_2} F_2(y) d(f_2)_*(\mathbb{P})(y)$$

by Fubini's Theorem. If both integrals exists, this is equal to

$$\int_{\Omega} F_1 \circ f_1 d\mathbb{P} \int_{\Omega} F_2 \circ f_2 d\mathbb{P}. \quad \square$$

Theorem 8.33. Let $B : [0, T] \times \Omega \rightarrow \mathbb{R}$ be a BM on \mathbb{R} . Fix $t_0 \in (0, T)$. Set $\tilde{B}_s(\omega) := B_{s+t_0}(\omega) - B_{t_0}(\omega)$. Then \tilde{B} and $B|_{[0, t_0]}$ are independent, i.e., $\tilde{B} : \Omega \rightarrow C_0([0, T - t_0]; \mathbb{R})$ is independent of $B|_{[0, t_0]} : \Omega \rightarrow C_0([0, t_0]; \mathbb{R})$. Also, both \tilde{B} and $B|_{[0, t_0]}$ are BMs on \mathbb{R} .

Proof. If $\sigma : [0, T] \rightarrow \mathbb{R}$ then define $\theta_{t_0}(\sigma) : [0, T - t_0] \rightarrow \mathbb{R}$ by $\theta_{t_0}(\sigma)(s) = \sigma(t_0 + s) - \sigma(t_0)$. The following diagram commutes:

$$\begin{array}{ccc} \Omega & \xrightarrow{\tilde{B} \times B|_{[0, t_0]}} & C_0([0, T - t_0]; \mathbb{R}) \times C_0([0, t_0]; \mathbb{R}) \\ & \searrow B & \nearrow \Psi_{t_0} \\ & & C_0([0, T]; \mathbb{R}) \end{array}$$

where Ψ_{t_0} is the product map $\Psi_{t_0}(\sigma) := (\theta_{t_0}(\sigma), \sigma|_{[0, t_0]})$. By Theorem 8.31, it is enough to know that Ψ_{t_0} sends Wiener measure to the product of Wiener measures. For this, see Exercises 3.4, 4.4, 2.6. \square

Corollary 8.34. *A BM has independent increments: fix $0 \leq s \leq t \leq u \leq v \leq T$. Then if B is a BM on \mathbb{R} , $B_t - B_s$ is independent of $B_v - B_u$, i.e.,*

$$B_t - B_s \perp B_v - B_u.$$

Proof. Set $\mathcal{F}_u = \mathcal{B}_u^B := \sigma\{B_r | 0 \leq r \leq u\}$ and $\mathcal{F}_t^u := \sigma\{B_{(u+r)} = B_u | 0 \leq r \leq T - u\}$. But $B_t - B_s$ is \mathcal{F}_u -measurable, and $B_v - B_u$ is \mathcal{F}_t^u -measurable. Therefore, $\sigma\{B_t - B_s\} \subseteq \mathcal{F}_u$ and $\sigma\{B_v - B_u\} \subseteq \mathcal{F}_t^u$, so $\sigma\{B_t - B_s\} \perp \sigma\{B_v - B_u\}$. \square

Theorem 8.35. *Given a probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$, $\mathcal{A} \subseteq \mathcal{F}$ a σ -algebra, and $f \in L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$, then*

$$f \perp \mathcal{A} \implies \mathbb{E}\{f | \mathcal{A}\} = \mathbb{E}f.$$

Proof. First, $\mathbb{E}f$ is \mathcal{A} -measurable and in L^1 . Secondly, if $A \in \mathcal{A}$, then

$$\int_A (\mathbb{E}f) d\mathbb{P} = (\mathbb{E}f) \int_A d\mathbb{P} = \mathbb{E}f \cdot \mathbb{E}\chi_A = \mathbb{E}(f \cdot \chi_A),$$

since $f \perp \chi_A$. By Corollary 8.32 with $F_1 = F_2 = \text{id}$, this is $\int_A f d\mathbb{P}$. Thus $\mathbb{E}f = \mathbb{E}\{f | \mathcal{A}\}$, as required. \square

Theorem 8.36. (The Martingale Property of Brownian Motion.) *If $\mathcal{F}_s = \mathcal{F}_s^B = \sigma\{B_r | 0 \leq r \leq s\}$ for a BM B on \mathbb{R} , then for $0 \leq s \leq t$,*

$$\mathbb{E}\{B_t | \mathcal{F}_s\} = B_s.$$

Proof. $B_t - B_s \perp \mathcal{F}_s$ by Corollary 8.34 and Theorem 8.35. Also,

$$\mathbb{E}\{B_t - B_s | \mathcal{F}_s\} = \mathbb{E}\{B_t - B_s\} = 0.$$

Since $B_r : \Omega \rightarrow \mathbb{R}$ is Gaussian, $\mathbb{E}B_r = 0$ for all r , and so $\int_{C_0} \sigma_\gamma d\gamma(\sigma) = 0$:

$$\begin{array}{ccc} \Omega, \mathbb{P} & \longrightarrow & C_0, \gamma \\ & \searrow B_r & \downarrow \text{ev} \\ & & \mathbb{R} \end{array} \quad \begin{array}{c} \sigma \\ \downarrow \\ \sigma_r \end{array}$$

But, since B_s is \mathcal{F}_s -measurable,

$$\begin{aligned} \mathbb{E}\{B_t - B_s | \mathcal{F}_s\} &= \mathbb{E}\{B_t | \mathcal{F}_s\} - \mathbb{E}\{B_s | \mathcal{F}_s\} \\ &= \mathbb{E}\{B_t | \mathcal{F}_s\} - B_s. \end{aligned} \quad \square$$

Theorem 8.37. If B is a BM on \mathbb{R} with its natural filtration $\mathcal{F}_s := \mathcal{F}_s^B$, then for $s \leq t$, $\mathbb{E}\{(B_t - B_s)^2 | \mathcal{F}_s\} = t - s$.

Proof. Theorem 8.36 implies that $B_t - B_s \perp \mathcal{F}_s$, so $(B_t - B_s)^2 \perp \mathcal{F}_s$, so

$$\begin{aligned} \mathbb{E}\{(B_t - B_s)^2 | \mathcal{F}_s\} &= \mathbb{E}\{(B_t - B_s)^2\} \\ &= \mathbb{E}\{B_t^2 - 2B_t B_s + B_s^2\} \\ &= t - 2(t \wedge s) + s \\ &= t - s. \end{aligned} \quad \square$$

Lemma 8.38. (Conditional Expectations.) Let $\{\Omega, \mathcal{F}, \mathbb{P}\}$ be a probability space, $\mathcal{A} \subseteq \mathcal{F}$ a σ -algebra, $\theta : \Omega \rightarrow \mathbb{R}$ \mathcal{A} -measurable and $f : \Omega \rightarrow \mathbb{R}$ \mathcal{F} -measurable.

(i) if $\theta, f \in L^2$ or θ bounded and $f \in L^1$,

$$\mathbb{E}\{\theta f | \mathcal{A}\} = \theta \mathbb{E}\{f | \mathcal{A}\};$$

(ii) if $\mathcal{B} \subseteq \mathcal{A}$ is a σ -algebra then

$$\mathbb{E}\{f | \mathcal{B}\} = \mathbb{E}\{\mathbb{E}\{f | \mathcal{A}\} | \mathcal{B}\}.$$

Proof. (i) For $\theta, f \in L^2$, write $\mathbb{E}\{f | \mathcal{A}\} = Pf$, where $P = P^{\mathcal{A}}$ is the orthogonal projection. Note that $\theta \mathbb{E}\{f | \mathcal{A}\}$ is \mathcal{A} -measurable and in L^1 . If $A \in \mathcal{A}$, then

$$\begin{aligned} \int_A \theta \cdot (f - Pf) \, d\mathbb{P} &= \int_{\Omega} \chi_A \theta (\text{id} - P)(f) \, d\mathbb{P} \\ &= \langle \chi_A \theta, (\text{id} - P)f \rangle_{L^2} \\ &= \langle (\text{id} - P)(\chi_A \theta), f \rangle_{L^2} \text{ since } (\text{id} - P)^* = (\text{id} - P) \\ &= 0 \text{ since } P(\chi_A \theta) = \chi_A \theta. \end{aligned}$$

Therefore, $\int_A \theta f \, d\mathbb{P} = \int_A \theta \mathbb{E}\{f | \mathcal{A}\} \, d\mathbb{P}$, as required.

Following from the above, if θ is bounded and \mathcal{A} -measurable, and $g \in L^1$, then $\mathbb{E}\{g\theta | \mathcal{A}\} = \theta \mathbb{E}\{g | \mathcal{A}\}$. If $f \in L^1$, take $f_n \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ with $f_n \rightarrow f$ in L^1 as $n \rightarrow \infty$. For instance, by the Dominated Convergence Theorem,

$$f_n(\omega) := \frac{f(\omega)}{1 + \frac{1}{n}|f(\omega)|}$$

will do. Again by the Dominated Convergence Theorem, $\theta f_n \rightarrow \theta f$ in L^1 , so

$$\begin{aligned} \theta \mathbb{E}\{f | \mathcal{A}\} &= \lim_{n \rightarrow \infty} \theta \mathbb{E}\{f_n | \mathcal{A}\} \\ &= \lim_{n \rightarrow \infty} \mathbb{E}\{\theta f_n | \mathcal{A}\} \\ &= \mathbb{E}\{\theta f | \mathcal{A}\} \end{aligned}$$

by the continuity of $\mathbb{E}\{- | \mathcal{A}\}$ in L^1 .

(ii) Easy exercise. □

Lemma 8.39. For B a BM on \mathbb{R} with its natural filtration $\mathcal{F}_* := \mathcal{F}_*^B$ and for a partition $0 \leq t_i < t_{i+1} \leq t_j < t_{j+1} \leq T$ of $[0, T]$, if $\alpha_i, \alpha_j : \Omega \rightarrow \mathbb{R}$ are bounded and \mathcal{F}_{t_i} - and \mathcal{F}_{t_j} -measurable respectively, then

$$(i) \quad \mathbb{E}\alpha_i\alpha_j\Delta_i B\Delta_j B = 0;$$

$$(ii) \quad \mathbb{E}\alpha_i^2(\Delta_i B)^2 = (\mathbb{E}\alpha_i^2)\Delta_i t.$$

Proof. (i) $\omega \mapsto \alpha_i(\omega)\alpha_j(\omega)(B_{t_{i+1}}(\omega) - B_{t_i})$ is \mathcal{F}_{t_j} -measurable.

$$\begin{aligned} \mathbb{E}\{\alpha_i\alpha_j\Delta_i B\Delta_j B\} &= \mathbb{E}\{\mathbb{E}\{\alpha_i\alpha_j\Delta_i B\Delta_j B|\mathcal{F}_{t_j}\}\} \\ &= \mathbb{E}\{\alpha_i\alpha_j\Delta_i B\mathbb{E}\{\Delta_j B|\mathcal{F}_{t_j}\}\} \text{ by Lemma 8.38} \\ &= 0 \end{aligned}$$

since $\mathbb{E}\{\Delta_j B|\mathcal{F}_{t_j}\} = 0$ (the martingale property).

(ii)

$$\begin{aligned} \mathbb{E}\alpha_i^2(\Delta_i B)^2 &= \mathbb{E}\{\mathbb{E}\{\alpha_i^2(\Delta_i B)^2|\mathcal{F}_{t_i}\}\} \\ &= \mathbb{E}\{\alpha_i^2\mathbb{E}\{(\Delta_i B)^2|\mathcal{F}_{t_i}\}\} \text{ by Lemma 8.38} \\ &= (\mathbb{E}\alpha_i^2)\Delta_i t \text{ by Lemma 8.38} \end{aligned} \quad \square$$

Remark 8.40. This proves Proposition 8.20, and so, by Theorem 8.21

$$\begin{aligned} \mathcal{I} : \mathcal{E}([0, T]) &\rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}) \\ a &\mapsto \int_0^T \alpha_j dB_j = \sum_j \alpha_j \Delta_j B \end{aligned}$$

is an isometry into $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ and so has a continuous linear extension

$$\bar{\mathcal{I}} : \bar{\mathcal{E}} \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$$

that is norm-preserving, where $\bar{\mathcal{E}}$ is the closure of \mathcal{E} in $L^2([0, T] \times \Omega, \mathcal{B}[0, T] \otimes \mathcal{F}, \lambda^1 \otimes \mathbb{P}; \mathbb{R})$. For $a \in \bar{\mathcal{E}}$ write $\bar{\mathcal{I}}(a)$ as $\int_0^T a_s dB_s$, the *Itô integral*. We usually write $L^2(B)$ for $\bar{\mathcal{E}}$, equipped with the norm

$$\|a\|_{L^2(B)} := \sqrt{\int_0^T \mathbb{E}(a_s)^2 ds}$$

$L^2(B)$ is also an inner product space: for $a, b \in L^2(B)$,

$$\langle a, b \rangle_{L^2(B)} := \int_0^T (\mathbb{E}a_s b_s) ds = \mathbb{E} \left\{ \int_0^T a_s dB_s \int_0^T b_s dB_s \right\}.$$

In particular, we have the *Itô isometry*:

$$\mathbb{E} \left(\int_0^T a_s dB_s \right)^2 = \mathbb{E} \left(\int_0^T (a_s)^2 ds \right).$$

Definitions 8.41. Given a filtration $\{\mathcal{F}_t | 0 \leq t \leq T\}$ of a probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$ and a measurable space $\{X, \mathcal{A}\}$, a process $a : [0, T] \times \Omega \rightarrow X$ is *progressively measurable* (or *progressive*) if for all $t \in [0, T]$ the map

$$\begin{aligned} [0, t] \times \Omega &\rightarrow X \\ (s, \omega) &\mapsto a_s(\omega) \end{aligned}$$

is $\mathcal{B}[0, t] \otimes \mathcal{F}_t$ -measurable (and so a is adapted). Also, we say that $P \subseteq [0, T] \times \Omega$ is *progressively measurable* if the process $a_s(\omega) := \chi_P(s, \omega)$ is progressive. The set of such P form a σ -algebra on $[0, T] \times \Omega$, denoted Prog , and a process a is progressive if, and only if, it is Prog -measurable.

It is a fact that $L^2(B)$ is the set of equivalence classes of Prog -measurable processes in $L^2([0, T] \times \Omega; \mathbb{R})$. Also, any adapted process with right- or left-continuous paths is Prog -measurable.

If a_s is independent of $\omega \in \Omega$ then we have both the Paley-Wiener integral $\int_0^T a_s dB_s$ and the Itô integral $\int_0^T a_s dB_s$, defined in different ways. Later work will show that they agree.

Definition 8.42. For $B : [0, T] \times \Omega \rightarrow \mathbb{R}$ a BM on \mathbb{R} , $a \in L^2(B)$ and $0 \leq t \leq T$, define

$$\int_0^t a_s dB_s := \int_0^T \chi_{[0, t]}(s) a_s dB_s.$$

Exercise 8.43. $a \in L^2(B) \implies \chi_{[0, t]} \cdot a \in L^2(B)$, so the RHS above makes sense.

Also note that $B|_{[0, t]}$ is a BM on \mathbb{R} and $a|_{[0, t]} \in L^2(B|_{[0, t]})$, so we can form $\int_0^t (a|_{[0, t]})_s d(B|_{[0, t]})_s$, and since it clearly agrees with $\int_0^t a_s dB_s$ for $a \in \mathcal{E}$, by continuity, it agrees for all $a \in L^2(B)$.

Definition 8.44. Let $\{\Omega, \mathcal{F}, \mathbb{P}\}$ be a probability space with filtration $\{\mathcal{F}_t | 0 \leq t \leq T\}$ or $\{\mathcal{F}_t | 0 \leq t < \infty\}$. A process $M : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ or $[0, \infty) \times \Omega \rightarrow \mathbb{R}^n$ is an \mathcal{F}_* -martingale if

- (i) it is adapted (i.e. $M_t : \Omega \rightarrow \mathbb{R}^n$ is \mathcal{F}_t -measurable);
- (ii) $M_t \in L^1$ for all t (and so $\mathbb{E}M_t$ exists for all t);
- (iii) the *martingale property*: if $s \leq t$ then $\mathbb{E}\{M_t | \mathcal{F}_s\} = M_s$ almost surely.

If no filtration is specified then M is a *martingale* if it is an \mathcal{F}_*^M -martingale.

Example 8.45. Brownian motions are martingales, as we have proved for $n = 1$.

Theorem 8.46. Let B be a BM on \mathbb{R} and $a \in L^2(B)$. Then the process $z : [0, T] \times \Omega \rightarrow \mathbb{R}$ defined by

$$z_t(\omega) := \left(\int_0^t a_s dB_s \right) (\omega)$$

is an \mathcal{F}_*^B -martingale.

Proof. If $a \in \mathcal{E}([0, T]; \mathbb{R})$, $s < t$,

$$a_r(\omega) = \alpha_{-1}(\omega) \chi_{\{0\}}(r) + \sum_{j=0}^k \alpha_j(\omega) \chi_{(t_j, t_{j+1}]}(r)$$

for α_{-1} \mathcal{F}_0 -measurable and α_j \mathcal{F}_{t_j} -measurable. We can assume that $s = t_{j_1}$, $t = t_{j_2}$ for some j_1, j_2 . Then

$$\int_0^t a_r dB_r = \int_0^s a_r dB_r + \sum_{j=j_1}^{j_2-1} \alpha_j \Delta_j B.$$

Now $\int_0^s a_r dB_r$ is \mathcal{F}_s -measurable and, applying the conditional expectation $\mathbb{E}\{-|\mathcal{F}_s\}$ to both sides,

$$\begin{aligned} \mathbb{E} \left\{ \sum_{j=j_1}^{j_2-1} \alpha_j \Delta_j B \middle| \mathcal{F}_s \right\} &= \sum_{j=j_1}^{j_2-1} \mathbb{E} \{ \mathbb{E} \{ \alpha_j \Delta_j B | \mathcal{F}_{t_j} \} | \mathcal{F}_s \} \\ &= \sum_{j=j_1}^{j_2-1} \mathbb{E} \{ \alpha_j \underbrace{\mathbb{E} \{ \Delta_j B | \mathcal{F}_{t_j} \}}_{(\bullet)} | \mathcal{F}_s \} \\ &= 0 \end{aligned}$$

since the martingale property of B implies that $(\bullet) = 0$. Therefore, for $a \in \mathcal{E}$, $\mathbb{E} \left\{ \int_0^t a_r dB_r \middle| \mathcal{F}_s \right\} = \int_0^s a_r dB_r$ almost surely. Since $\mathbb{E}\{-|\mathcal{F}_s\}$ is continuous in L^2 , the result follows for $a \in L^2(B)$. \square

Corollary 8.47. *Let B be a BM on \mathbb{R} and let $a \in L^2(B)$. Then $\mathbb{E} \int_0^t a_s dB_s = 0$ for all $t \in [0, T]$.*

Proof. Since $z_t := \int_0^t a_s dB_s$ is an \mathcal{F}_*^B -martingale,

$$\mathbb{E} z_t = \mathbb{E} \{ z_t | \mathcal{F}_0^B \} = z_0 = 0,$$

and $z_t \perp \mathcal{F}_0^B$ since $\mathcal{F}_0^B = \{\emptyset, \Omega\}$. \square

Remark 8.48. It can be proved that $[0, T] \rightarrow \mathbb{R} : t \mapsto \int_0^t a_s dB_s(\omega)$ may be chosen to be continuous (so that $\int_0^t a_s dB_s$ for $0 \leq t \leq T$ is sample continuous). For each t we have to choose some version of $\int_0^t a_s dB_s$ from its L^2 -equivalence class. But if we want $\int_0^t a_s dB_s$ to be \mathcal{F}_t^B -measurable as well, we need to modify \mathcal{F}_t^B to include sets of measure zero in \mathcal{F}_T^B . Then we get a process that is both continuous and adapted. These are the “usual conditions” on \mathcal{F}_* .

Theorem 8.49. (Itô’s Formula.) *Let B be a BM on \mathbb{R} with its natural filtration $\mathcal{F}_t = \mathcal{F}_t^B$. Let $z : [0, T] \times \Omega \rightarrow \mathbb{R}$ be given by*

$$z_t(\omega) = a_t(\omega) + \left(\int_0^t \alpha_s dB_s \right) (\omega),$$

for an adapted $a : [0, T] \times \Omega \rightarrow \mathbb{R}$ such that $t \mapsto a_t(\omega)$ is piecewise C^1 (or of bounded variation), and $\alpha \in L^2(B)$. Suppose that $\theta : \mathbb{R} \rightarrow \mathbb{R}$ is C^2 . Then for $0 \leq t \leq T$,

$$\begin{aligned} \theta(z_t(\omega)) &= \theta(z_0(\omega)) + \int_0^t \theta'(z_s(\omega)) a'_s(\omega) ds + \left(\int_0^t \theta'(z_s(-)) \alpha_s dB_s \right) (\omega) \\ &\quad + \frac{1}{2} \int_0^t \theta''(z_s(\omega)) \alpha_s(\omega) \alpha_s(\omega) ds \text{ almost surely.} \end{aligned}$$

For us, we need $\theta'(z_s(\omega))\alpha_s(\omega)$ in $L^2(B)$. The idea of the proof is to use stopping times to extend our definition of the integrand. We take $0 = t_0 < t_1 < \dots < t_{k+1} = t$. For “nice” θ ,

$$\begin{aligned} & \theta(z_t(\omega)) - \theta(z_0(\omega)) \\ &= \sum_{j=0}^k (\theta(z_{t_{j+1}}(\omega)) - \theta(z_{t_j}(\omega))) \\ &= \sum_{j=0}^k \left(\theta'(z_{t_j}(\omega)) (z_{t_{j+1}}(\omega) - z_{t_j}(\omega)) + \frac{1}{2} \theta''(z_{t_j}(\omega)) (z_{t_{j+1}}(\omega) - z_{t_j}(\omega))^2 + \text{higher order} \right) \end{aligned}$$

The first and second terms in Itô’s formula come from the first term here.

$$\Delta_j z \approx (a'_{t_j}) \Delta_j t + \alpha_{t_j} \Delta_j B$$

so

$$\begin{aligned} (\Delta_j z)^2 &\approx (a'_{t_j})^2 (\Delta_j t)^2 + 2a'_{t_j} \alpha_{t_j} \Delta_j B \Delta_j t + (\alpha_{t_j})^2 (\Delta_j B)^2 \\ &\approx 0 + 0 + (\alpha_{t_j})^2 \Delta_j t \end{aligned}$$

as we know that $\mathbb{E}\{(\Delta_j B)^2\} = \Delta_j t$. We summarize these results in the Itô multiplication table:

	dt	dB
dt	0	0
dB	0	dt

See exercises for a more in-depth treatment of this.

Examples 8.50. (i) $z_t = B_t = 0 + \int_0^t dB_s$; this is $a \equiv 0$, $\alpha \equiv 1$ in Itô’s formula. Let $\theta : \mathbb{R} \rightarrow \mathbb{R}$ be $\theta(x) = x^2$. So

$$\begin{aligned} B_t^2 &= B_0^2 + \int_0^t 2B_s dB_s + \frac{1}{2} \int_0^t 2 ds \\ &= 0 + 2 \int_0^t B_s dB_s + t. \end{aligned}$$

Thus,

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

(ii) Exponential martingales. Let $h \in L_0^{2,1}([0, T]; \mathbb{R})$. Since $\dot{h} \in L^2$, $h(t) = \int_0^t \dot{h}(s) ds$. For $0 \leq t \leq T$, set

$$M_t = \exp \left(\int_0^t \dot{h}(s) dB_s - \frac{1}{2} \int_0^t |\dot{h}(s)|^2 ds \right),$$

using either the Paley-Wiener or Itô integral as h is independent of ω .

Claim 8.51. M satisfies

$$M_t = 1 + \int_0^t M_s \dot{h}(s) dB_s$$

almost surely for $0 \leq t \leq T$.

This is an example of a *stochastic differential equation*:

$$\begin{cases} dM_t = M_t \dot{h}_t dB_t \\ M_0 = 1 \end{cases}$$

Proof. Set

$$z_t = \int_0^t \dot{h}(s) dB_s - \frac{1}{2} \int_0^t |\dot{h}(s)|^2 ds$$

with $\theta : \mathbb{R} \rightarrow \mathbb{R}$ given by $\theta(x) = e^x$. Then

$$\begin{aligned} \theta(z_t(\omega)) &= \theta(z_t(\omega)) + \int_0^t e^{z_s(\omega)} \left(-\frac{1}{2} \dot{h}_s^2 \right) ds \\ &\quad + \left(\int_0^t e^{z_s(\omega)} \dot{h}_s dB_s \right) (\omega) + \frac{1}{2} \int_0^t e^{z_s(\omega)} (\dot{h}_s)^2 ds \end{aligned}$$

almost surely for $0 \leq t \leq T$. So

$$M_t = 1 + \int_0^t M_s \dot{h}(s) dB_s \text{ almost surely.} \quad \square$$

We should check that $M_s \dot{h}(s) \in L^2(B_s)$, i.e., $M \cdot \dot{h}(\cdot) \in L^2(B)$.

Lemma 8.52. $M_t \in L^p$ for $1 \leq p < \infty$. In fact,

$$\mathbb{E}\{(M_t)^p\} = e^{\frac{1}{2}p(p-1) \int_0^t (h_s)^2 ds}.$$

Proof. Method 1. By Proposition 7.1, with $H = L_0^{2,1}$, for $w \in \mathbb{C}$,

$$\int_{C_0} e^{w \langle h, - \rangle_{\tilde{H}}(\sigma)} d\gamma(\sigma) = e^{\frac{1}{2}w^2 \|h\|_H^2}.$$

Therefore,

$$\begin{aligned} \int_{C_0} e^{p \langle h, - \rangle_{\tilde{H}} - \frac{1}{2}p \|h\|_H^2} d\gamma &= e^{\frac{1}{2}p^2 \|h\|_H^2 - \frac{1}{2}p \|h\|_H^2} \\ &= e^{\frac{1}{2}p(p-1) \|h\|_H^2} \end{aligned}$$

But M_t is the composition

$$\Omega \xrightarrow{B|_{[0,t](-)}} C_0([0,t]; \mathbb{R}) \xrightarrow{e^{\langle h, - \rangle_{\tilde{H}} - \frac{1}{2} \|h\|_H^2}} \mathbb{R}.$$

Method 2. Use Cameron-Martin: in the Wiener space $C_0([0,t]; \mathbb{R})$, if $F : C_0 \rightarrow \mathbb{R}$ is measurable, then for $p \in \mathbb{R}$,

$$\int_{C_0} F(\sigma + ph) d\gamma(\sigma) = \int_{C_0} F(\sigma) e^{p \langle h, - \rangle_{\tilde{H}} - \frac{1}{2}p^2 \|h\|_H^2} d\gamma(\sigma),$$

so

$$\mathbb{E}\{F(B + ph)\} = \mathbb{E}\left\{F(B)(M_t)^p e^{\frac{1}{2}p \|h\|_H^2 - \frac{1}{2}p^2 \|h\|_H^2}\right\}.$$

Now take $F \equiv 1$ to get

$$\begin{aligned} 1 &= \int_{C_0} e^{p \langle h, - \rangle_{\tilde{H}} - \frac{1}{2}p^2 \|h\|_H^2} \\ &= \mathbb{E}\{(M_t)^p\} e^{\frac{1}{2}p \|h\|_H^2 - \frac{1}{2}p^2 \|h\|_H^2}. \end{aligned} \quad \square$$

Proposition 8.53. M_t , $0 \leq t \leq T$, is an \mathcal{F}_*^B -martingale. In particular, $\mathbb{E}M_t = 1$ for all $0 \leq t \leq T$.

Proof. This is immediate from Claim 8.51. □

Remark 8.54. As was noted above, an exponential martingale is an example of a stochastic differential equation. A general stochastic differential equation on R takes the form

$$\begin{cases} dx_t = A(x_t) dt + X(x_t) dB_t \\ x_0 = q. \end{cases}$$

This means that $x : [0, T] \times \Omega \rightarrow \mathbb{R}$ satisfies

$$x_t = q + \int_0^t A(x_s) ds + \int_0^t X(x_s) dB_s \text{ almost surely.}$$

q could be a point of \mathbb{R} or a function $q : \Omega \rightarrow \mathbb{R}$; we need x to be adapted. Note that if $X(x) = 0$ for all x , we have an ordinary differential equation

$$\frac{dx_t}{dt} = A(x_t).$$

9 Itô Integrals as Divergences

9.1 The Clark-Ocone Theorem and Integral Representation

We work with canonical 1-dimensional Brownian motion:

$$\begin{aligned}\Omega &= C_0([0, T]; \mathbb{R}) \\ \mathcal{F}_t &= \sigma\{\text{ev}_s \mid 0 \leq s \leq t\} = \sigma\{\rho \mapsto \rho(s) \mid 0 \leq s \leq t\} \\ B_t(\omega) &= \omega(t) = \text{ev}_t(\omega)\end{aligned}$$

so $\mathcal{F}_t = \mathcal{F}_t^B$.

Theorem 9.1. *For $V : [0, T] \times C_0 \rightarrow \mathbb{R}$ such that*

$$V_t(\sigma) = \sum_{j=0}^k (t \wedge t_{j+1} - t \wedge t_j) \alpha_j(\sigma)$$

where $0 = t_0 < t_1 < \dots < t_{k+1} = T$ and $\alpha_j : \Omega \rightarrow \mathbb{R}$ has $\alpha_j(\sigma)$ depending only on $\sigma|_{[0, t_j]}$, i.e. α_j is \mathcal{F}_{t_j} -measurable, and α_j is bounded for all j , then if $F : C_0 \rightarrow \mathbb{R}$ is measurable, then

$$\int_{C_0} F(\sigma + V(\sigma)) e^{-\int_0^T \frac{\partial}{\partial t} V_t(\sigma) d\sigma(t) + \frac{1}{2} \int_0^T \left| \frac{\partial}{\partial t} V_t(\sigma) \right|^2 dt} d\gamma(\sigma) = \int_{C_0} F(\sigma) d\gamma(\sigma),$$

where γ is Wiener measure.

Proof. We use induction on k . Consider the case $k = 0$. $0 = t_0 < t_1 = T$ and α_0 is constant, since it is \mathcal{F}_0 -measurable and $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Thus $V_t = t\alpha_0$ for $0 \leq t \leq T$, so $V \in H := L_0^{2,1}([0, T]; \mathbb{R})$ and we can apply the Cameron-Martin Formula, Theorem 7.7, with $\check{F} : C_0 \rightarrow \mathbb{R}$ given by $\check{F}(\sigma) := F(\sigma + V)$. Thus

$$\int_{C_0} \check{F}(\sigma + V) d\gamma(\sigma) = \int_{C_0} \check{F}(\sigma) e^{-(V, \cdot) \sim (\sigma) - \frac{1}{2} \|V\|_{L^2}^2} d\gamma(\sigma),$$

i.e.

$$\int_{C_0} F(\sigma) d\gamma(\sigma) = \int_{C_0} F(\sigma + V(\sigma)) e^{-\int_0^T \dot{V}_s d\sigma_s - \frac{1}{2} \int_0^T |\dot{V}_s|^2 ds} d\gamma(\sigma)$$

Now assume true for $k = n - 1$ for some $n \in \mathbb{N}$ and consider the case $k = n$. Set $T_0 = t_n$ so $0 = t_0 < \dots < t_n = T_0 < t_{n+1} = T$. We have

$$\begin{array}{ccc} C_0([0, T]; \mathbb{R}) & \xrightarrow{\Theta} & C_0([0, T_0]; \mathbb{R}) \times C_0([0, T - T_0]; \mathbb{R}) \\ T_V \downarrow & & \downarrow \tilde{T}_V \\ C_0([0, T]; \mathbb{R}) & \xrightarrow{\Theta} & C_0([0, T_0]; \mathbb{R}) \times C_0([0, T - T_0]; \mathbb{R}) \end{array}$$

where $T_V(\sigma) := \sigma + V(\sigma)$, $\tilde{T}_V(\sigma, \rho) := \Theta(\tilde{\sigma} + V(\tilde{\sigma}))$, where

$$\tilde{\sigma}(t) := \begin{cases} \sigma(t) & 0 \leq t \leq T_0 \\ \sigma(T_0) + \rho(t - T_0) & T_0 \leq t \leq T, \end{cases}$$

so $\tilde{T}_V = \Theta \circ T_V \circ \Theta^{-1}$, $\tilde{\sigma} = \Theta^{-1}(\sigma, \rho)$. Then

$$V_t(\tilde{\sigma}) = \begin{cases} \sum_{j=0}^{k-1} (t \wedge t_{j+1} - t \wedge t_j) \alpha_j(\sigma) =: V_t^{T_0}(\sigma) & 0 \leq t \leq T_0 \\ (t - T_0) \alpha_k(\sigma) & T_0 \leq t \leq T, \end{cases}$$

since $\alpha_j(\tilde{\sigma})$ depends only on $\tilde{\sigma}|_{[0, t_j]}$, so $\alpha_j(\tilde{\sigma}) = \alpha_j(\sigma)$. Thus, $\tilde{T}_V(\sigma, \rho) = (\sigma + V^{T_0}(\sigma), \rho + t \alpha_k(\sigma))$. Also,

$$\int_0^T \frac{\partial}{\partial s} V_s(\tilde{\sigma}) d\tilde{\sigma}(s) = \int_0^{T_0} \frac{\partial}{\partial s} V_s(\sigma) d\tilde{\sigma}(s) + \int_0^{T-T_0} \alpha_k(\sigma) d\rho(s),$$

since these are elementary Itô integrals and so just a sum of $\Delta_j \sigma_s$ because $\frac{\partial}{\partial s} V_s \in \mathcal{E}$. Set

$$\tilde{F} := F \circ \Theta^{-1} : C_0([0, T_0]; \mathbb{R}) \times C_0([0, T - T_0]; \mathbb{R}) \rightarrow \mathbb{R}.$$

Then

$$\begin{aligned} & \int_{C_0} F(\sigma + V(\sigma)) e^{-\int_0^T \frac{\partial}{\partial s} V_s d\sigma_s - \frac{1}{2} \int_0^T \left(\frac{\partial}{\partial s} V_s\right)^2 ds} d\gamma(\sigma) \\ &= \int_{C_0^{T_0} \times C_0^{T-T_0}} \tilde{F} \circ \tilde{T}_V e^{-\int_0^{T_0} \dot{V}_s d\sigma_s - \frac{1}{2} \int_0^{T_0} (\dot{V}_s(\sigma))^2 ds} e^{-\int_0^{T-T_0} \dot{\alpha}(\sigma)_s d\rho_s - \frac{1}{2} \int_0^{T-T_0} (\dot{\alpha}_k(\sigma)_s)^2 ds} d\gamma^{T-T_0}(\rho) d\gamma^{T_0}(\sigma) \end{aligned}$$

where γ^τ denotes Wiener measure on $C_0^\tau := C_0([0, \tau]; \mathbb{R})$, $\tau \in \{T_0, T - T_0\}$

$$= \int_{C_0^{T_0}} \left\{ \int_{C_0^{T-T_0}} \tilde{F}(\sigma + V^{T_0}(\sigma), \rho) d\gamma^{T-T_0}(\rho) \right\} e^{-\int_0^{T_0} \dot{V}_s^{T_0} d\sigma_s - \frac{1}{2} \int_0^{T_0} (\dot{V}_s^{T_0})^2 ds} d\gamma^{T_0}(\sigma)$$

by applying the Cameron-Martin Formula to γ^{T-T_0} as in the case $k = 0$, and so, by the induction hypothesis,

$$\begin{aligned} &= \int_{C_0^{T_0}} \left\{ \int_{C_0^{T-T_0}} \tilde{F}(\sigma, \rho) d\gamma^{T-T_0}(\rho) \right\} d\gamma^{T_0}(\sigma) \\ &= \int_{C_0} F(\sigma) d\gamma(\sigma) \end{aligned}$$

as required. \square

Lemma 9.2. (Improved Integration by Parts.) *For V as in Theorem 9.1, $F : C_0 \rightarrow \mathbb{R}$ of class¹ BC^1 and $\gamma =$ Wiener measure,*

$$\int_{C_0} DF(\sigma)(V(\sigma)) d\gamma(\sigma) = \int_{C_0} F(\sigma) \left(\int_0^T \frac{\partial V}{\partial s}(\sigma) d\sigma(s) \right) d\gamma(\sigma).$$

Proof. For $\tau \in \mathbb{R}$ replace V by τV in the formula of Theorem 9.1 and differentiate both sides with respect to τ and evaluate at $\tau = 0$. \square

Definition 9.3. If $i : H \rightarrow E$ is an AWS (such as $L_0^{2,1} \rightarrow C_0$) and $F : E \rightarrow \mathbb{R}$ is differentiable we get $DF(x) \in \mathbb{L}(E; \mathbb{R}) = E^*$ for $x \in E$. So, for all $x \in E$, $D_H F(x) := DF(x) \circ i : H \rightarrow \mathbb{R}$ is a continuous linear map, the *derivative of F in H -direction* or *H -derivative*.

¹ F is bounded and Fréchet differentiable with $DF : C_0 \rightarrow \mathbb{L}(C_0; \mathbb{R})$ bounded.

Thus, we get $\nabla_H F : E \rightarrow H$ defined by

$$\langle \nabla_H F(x), h \rangle_H = D_H F(x)(h) = \lim_{t \rightarrow 0} \frac{F(x + ti(h)) - F(x)}{t}.$$

So $\nabla_H F(x) = j(DF(x))$, with $j : E^* \rightarrow H$ as usual. Therefore, if F is C^1 , then $\nabla_H F : E \rightarrow H$ is continuous. If F is BC^1 then $\|\nabla_H F(x)\| \leq \|j\| \|DF(x)\|_{E^*} \leq \text{constant}$, so $\nabla_H F$ is bounded.

Remark 9.4. Lemma 9.2 can be written

$$\int_{C_0} \langle \nabla_H F(x), V(x) \rangle_H d\gamma(x) = - \int_{C_0} F(x) \operatorname{div} V(x) d\gamma(x)$$

where $\operatorname{div} V : C_0 \rightarrow \mathbb{R}$ is $-\int_0^T \dot{V}_s dB_s$.

Definition 9.5. If E is a normed vector space, a subset $S \subseteq E$ is *total* in E if the span of S is dense in E :

$$\operatorname{span} S := \left\{ \sum_{j=1}^k \alpha_j x_j \mid \alpha_j \in \mathbb{R}, x_j \in S, k \in \mathbb{N} \right\}.$$

Lemma 9.6. S is total in a Hilbert space H if, and only if,

$$\langle h, s \rangle_H = 0 \forall s \in S \implies h = 0.$$

Proof. $\langle h, s \rangle_H = 0 \forall s \in S \iff h \perp \operatorname{span} S \iff h \perp \overline{\operatorname{span} S}$ and $\overline{\operatorname{span} S^\perp} = 0 \iff \overline{\operatorname{span} S} = H$. \square

Proposition 9.7. In an AWS $i : H \rightarrow E$, $\left\{ e^{\langle h, - \rangle_{\tilde{H}} - \frac{1}{2} \|h\|_H^2} \mid h \in H \right\}$ is total in $L^2(E, \gamma; \mathbb{R})$. (In fact, we need only that $h \in j(E^*)$.)

Proof. Note that $e^{-\frac{1}{2} \|h\|_H^2}$ is constant and so is irrelevant. Suppose $f : E \rightarrow \mathbb{R}$ is in L^2 with $\int_E f(x) e^{\langle h, - \rangle_{\tilde{H}}(x)} d\gamma(x) = 0$ for all $h \in j(E^*)$. Taking $h = j(\ell)$ this gives that for all $\ell \in E^*$,

$$\int_E f(x) e^{\ell(x)} d\gamma(x) = 0.$$

Note that $z \mapsto \int_E e^{z\ell(x)} f(x) d\gamma(x) : \mathbb{C} \rightarrow \mathbb{C}$ is analytic in $z \in \mathbb{C}$, as usual. Therefore, for all $z \in \mathbb{C}$ and $\ell \in E^*$,

$$\int_E e^{z\ell(x)} f(x) d\gamma(x) = 0,$$

and so, for all $\ell \in E^*$,

$$\int_E e^{i\ell(x)} f(x) d\gamma(x) = 0,$$

The result then follows from Lemma 9.8. \square

Lemma 9.8. If $f \in L^1(E, \mu; \mathbb{R})$ is such that

$$\int_E f(x) e^{i\ell(x)} d\mu(x) = 0$$

for all $\ell \in E^*$, where μ is a finite measure on E , then $f = 0$ μ -almost surely.

Proof. Set $f = f^+ - f^-$ as usual, so $f^+(x)f^-(x) = 0$ for all $x \in E$. Form measures μ_{f^\pm} by $\mu_{f^\pm}(A) := \int_A f^\pm(x) d\mu(x)$. Then $\int_E f(x)e^{i\ell(x)} d\mu(x) = 0$ for all $\ell \in E^*$, so $\widehat{\mu_{f^+}} = \widehat{\mu_{f^-}}$, and so $\mu_{f^+} = \mu_{f^-}$ by Bochner's Theorem. Thus,

$$\begin{aligned} \int_E (f^+(x))^2 d\mu(x) &= \int_E f^+(x) d\mu_{f^+}(x) \\ &= \int_E f^+(x) d\mu_{f^-}(x) \\ &= \int_E f^+(x)f^-(x) d\mu(x) \\ &= 0. \end{aligned}$$

So $f^+ = 0$ almost surely, as does f^- , and, therefore, so does f . \square

Remark 9.9. Lemma 9.8 shows that $\{\sin \ell(\cdot), \cos \ell(\cdot) | \ell \in E^*\}$ is total in $L^2(E, \mu; \mathbb{R})$, since $\cos \ell(x) + i \sin \ell(x) = e^{i\ell(x)}$, so $f \perp \sin \ell(\cdot), f \perp \cos \ell(\cdot) \implies f = 0$ by Lemma 9.8. In particular, the BC^1 functions $E \rightarrow \mathbb{R}$ are dense in $L^2(E, \mu; \mathbb{R})$.

Proposition 9.10. $\left\{1, \sigma \mapsto \int_0^T \alpha_s d\sigma(s) \mid \alpha \in \mathcal{E}\right\}$ is total in $L^2(C_0, \gamma; \mathbb{R})$ for $\gamma =$ Wiener measure on $C_0([0, T]; \mathbb{R})$.

Proof. For $h \in H$ and $0 \leq t \leq T$, set

$$M_t^h := \exp\left(\int_0^t \dot{h}_s d\sigma(s) + \frac{1}{2} \int_0^t |\dot{h}_s|^2 ds\right).$$

By Claim 8.51, $M_T^h = 1 + \int_0^T M_s^h \dot{h}_s d\sigma(s)$. Therefore, $\left\{1, \sigma \mapsto \int_0^T \alpha_s d\sigma(s) \mid \alpha \in L^2(B)\right\}$ is total in $L^2(C_0; \mathbb{R})$ by Proposition 9.7. But each $\int_0^T \alpha_s d\sigma(s)$ is an L^2 limit as $n \rightarrow \infty$ of a sequence $\int_0^T \alpha_s^n d\sigma(s)$ for some $\alpha^n \in \mathcal{E}$. \square

Theorem 9.11. (Clark-Ocone Theorem for BC^1 Functions.) *If $F : C_0 \rightarrow \mathbb{R}$ is BC^1 then*

$$F(\sigma) = \int_{C_0} F d\gamma + \int_0^T \mathbb{E} \left\{ \frac{\partial}{\partial t} \nabla_H F_t(\cdot) \mid \mathcal{F}_t \right\}(\sigma) d\sigma(t),$$

where \mathcal{F}_t is the natural filtration of canonical BM, $\sigma\{\text{ev}_s \mid 0 \leq s \leq t\}$.

Proof. Set $G(\sigma) := \int_0^T \dot{V}_s(\sigma) d\sigma(s)$; \dot{V} is elementary, so V is as in Theorem 9.1. By Proposition 9.10, the set of such G together with the constants is total in L^2 . Set $\bar{F} := F - \int_{C_0} F d\gamma$, so $\int_{C_0} \bar{F} d\gamma = 0$, so $\bar{F} \perp$ all constants in L^2 . Then

$$\begin{aligned} \int_{C_0} \bar{F}(\sigma) G(\sigma) d\gamma(\sigma) &= \int_{C_0} \langle \nabla_H F(\sigma), V(\sigma) \rangle_{L_0^{2,1}} d\gamma(\sigma) \\ &= \int_{C_0} \left\{ \int_0^T \dot{\nabla}_H F_s(\sigma) \dot{V}_s(\sigma) ds \right\} d\gamma(\sigma) \end{aligned}$$

Now $\dot{\nabla}_H F(\sigma) \in L^2([0, T]; \mathbb{R})$, bounded in $\sigma \in C_0$, and so is $\dot{V}(\sigma)$. Therefore, $\dot{\nabla}_H F(\sigma)\dot{V}(\sigma) \in L^1([0, T] \times \Omega; \mathbb{R})$, so we can apply Fubini's Theorem to the above:

$$\begin{aligned} \text{RHS} &= \int_0^T \left(\int_{C_0} \nabla_H F_s(\sigma) \dot{V}_s(\sigma) d\gamma(\sigma) \right) ds \\ &= \int_0^T \left(\int_{C_0} \mathbb{E} \left\{ \nabla_H F_s(\sigma) \middle| \mathcal{F}_s \right\} \dot{V}_s(\sigma) d\gamma(\sigma) \right) ds \end{aligned}$$

since $\dot{V}_s(\sigma)$ is \mathcal{F}_s -measurable

$$\begin{aligned} &= \int_{C_0} \left(\int_0^T \mathbb{E} \left\{ \dot{\nabla}_H F_s(-) \middle| \mathcal{F}_s \right\}(\sigma) d\sigma(s) \right) \left(\int_0^T \dot{V}_s(\sigma) d\sigma(s) \right) d\gamma(\sigma) \\ &= \int_{C_0} \left(G(\sigma) \int_0^T \mathbb{E} \left\{ \dot{\nabla}_H F_s(-) \middle| \mathcal{F}_s \right\}(\sigma) d\sigma(s) \right) d\gamma(\sigma) \end{aligned}$$

by the isometry property of the Itô integral. Now subtract the LHS from the RHS and use the totality of the G s and constants together with the fact that the expectation of an Itô integral is zero. Therefore,

$$\left(\bar{F}(\sigma) - \int_0^T \mathbb{E} \left\{ \nabla_H F_s(\sigma) \middle| \mathcal{F}_s \right\} d\sigma(s) \right) \perp G$$

for all G . Since

$$\int_{C_0} \left(\bar{F}(\sigma) - \int_0^T \mathbb{E} \left\{ \nabla_H F_s(-) \middle| \mathcal{F}_s \right\}(\sigma) d\sigma(s) \right) d\gamma(\sigma) = 0,$$

it is orthogonal to all constants. Thus, by the totality of $\{G, \text{constants}\}$,

$$\bar{F} = \int_0^T \mathbb{E} \left\{ \nabla_H F_s(-) \middle| \mathcal{F}_s \right\} d\sigma(s),$$

as claimed. □

Remark 9.12. This says that for $F \in BC^1$, $F = \int F + \text{div} U$ for some U , since if F is BC^1 , $F = \text{div} U$ for nice $U \iff \int_{C_0} F = 0$. The result for $F \in L^2$ is called the Integral Representation Theorem.

Theorem 9.13. (Integration by Parts on C_0 .) *Let $V : C_0 \rightarrow L_0^{2,1}$ be such that $\dot{V} : [0, T] \times C_0 \rightarrow \mathbb{R}$ is in $L^2(B)$, and so is adapted. Let $F : C_0 \rightarrow \mathbb{R}$ be BC^1 . Then*

$$\int_{C_0} DF(\sigma)(V(\sigma)) d\gamma(\sigma) = \int_{C_0} F(\sigma) \left(\int_0^T \dot{V}_s(\sigma) d\sigma(s) \right) d\gamma(\sigma)$$

i.e.

$$\int_{C_0} \langle \nabla_H F(\sigma), V(\sigma) \rangle_{L_0^{2,1}} d\gamma(\sigma) = - \int_{C_0} F(\sigma) \text{div} V(\sigma) d\gamma(\sigma),$$

where $\text{div} V : C_0 \rightarrow \mathbb{R}$ is $\text{div} V(\sigma) := - \int_0^T \dot{V}_s(\sigma) d\sigma(s)$.

Proof. Use Clark-Ocone to substitute for F in the RHS since, by the martingale property of Itô integrals,

$$\int_{C_0} \left(\int_0^T \dot{V}(\sigma)_s d\sigma(s) \right) d\gamma(\sigma) = 0.$$

Thus

$$\begin{aligned} \text{RHS} &= 0 + \int_{C_0} \left(\int_0^T \mathbb{E} \left\{ \nabla_H \dot{F}(\sigma)_s \mid \mathcal{F}_s \right\} \int_0^T \dot{V}(\sigma)_s d\sigma(s) \right) d\gamma(\sigma) \\ &= \int_0^T \left(\int_{C_0} \mathbb{E} \left\{ \nabla_H \dot{F}(\sigma)_s \mid \mathcal{F}_s \right\} \dot{V}(\sigma)_s \right) d\gamma(\sigma) ds \\ &= \int_0^T \int_{C_0} \nabla_H \dot{F}(\sigma)_s \dot{V}(\sigma)_s d\gamma(\sigma) ds \text{ since } \dot{V}(-) \text{ is } \mathcal{F}_s\text{-measurable} \\ &= \int_{C_0} \langle \nabla_H F(\sigma), V(\sigma) \rangle_{L_0^{2,1}} d\gamma(\sigma) \\ &= \text{LHS} \end{aligned} \quad \square$$

The above result shows that Itô integrals are divergences.

Theorem 9.14. (Integral Representation for C_0 .) *If $F \in L^2(C_0; \mathbb{R})$ then there exists a unique $\alpha^F : [0, T] \times C_0 \rightarrow \mathbb{R}$ in $L^2(B)$ such that*

$$F(\sigma) = \int_{C_0} F d\gamma + \int_0^T \alpha^F(\sigma)_s d\sigma(s) \text{ almost surely.}$$

Proof. For an alternative proof, see [MX]. To show existence, set $\hat{L}^2 := \{f \in L^2 \mid \mathbb{E}f = 0\}$. If $F \in \hat{L}^2 \cap BC^1$ then

$$F(\sigma) = \int_0^T U(F)(\sigma)_s d\sigma(s),$$

where $U(F)_s = \mathbb{E} \left\{ \frac{\partial}{\partial t} \nabla_H F(-)_s \mid \mathcal{F}_s \right\} \in L^2(B)$,

$$U : \hat{L}^2 \cap BC^1 \rightarrow L^2(B) \subset L^2([0, T] \times C_0; \mathbb{R}).$$

Recall the the Itô integral $\bar{\mathcal{J}} : L^2(B) \rightarrow L^2$ is norm-preserving. Thus, Clark-Ocone implies that $\bar{\mathcal{J}} \circ U(F) = F$ for all $F \in \hat{L}^2 \cap BC^1$, and

$$\|F\|_{L^2} = \|\bar{\mathcal{J}} \circ U(F)\|_{L^2} = \|U(F)\|_{L^2(B)},$$

so U preserves the L^2 norm. But BC^1 is dense in L^2 by Remark 9.9, so $\hat{L}^2 \cap BC^1$ is dense in \hat{L}^2 , since the projection $F \mapsto F - \int_{C_0} F d\gamma$ maps BC^1 to $\hat{L}^2 \cap BC^1$. Therefore, U has a unique continuous linear extension $\bar{U} : \hat{L}^2 \rightarrow L^2(B)$ that is norm-preserving. Since $\bar{\mathcal{J}} \circ U = \text{id}$ on the dense subset $\hat{L}^2 \cap BC^1$, $\bar{\mathcal{J}} \circ \bar{U} = \text{id}$ on \hat{L}^2 . So for $f \in \hat{L}^2$,

$$f = \int_0^T \bar{U}(f_s)(\sigma) d\sigma(s).$$

So for $F \in L^2$, set

$$\alpha^F := \bar{U} \left(F - \int_{C_0} F d\gamma \right).$$

As for uniqueness, suppose we are given two candidates α^F and $\tilde{\alpha}^F$. Set $\beta := \alpha^F - \tilde{\alpha}^F \in L^2(B)$. Then $\int_0^T \beta(\sigma)_s d\sigma(s) = 0$. Therefore,

$$\sqrt{\int_0^T \|\beta(\sigma)_s\|_{L^2}^2 ds} = \left\| \int_0^T \beta(\sigma)_s d\sigma_s \right\|_{L^2} = 0,$$

so $\|\beta\|_{L^2(B)} = 0$ by the isometry property, i.e. $\beta = 0$. \square

Theorem 9.15. (Integral Representation.) *Let $\{\Omega, \mathcal{F}, \mathbb{P}\}$ be a probability space, $B : [0, T] \times \Omega \rightarrow \mathbb{R}$ a BM with filtration $\mathcal{F}_* = \mathcal{F}_*^B$. Suppose $f : \Omega \rightarrow \mathbb{R}$ is in L^2 and \mathcal{F}_T -measurable. Then there is a unique $a^f \in L^2(B)$ such that*

$$f = \mathbb{E}f + \int_0^T a_s^f dB_s \text{ almost surely.}$$

Proof. Consider the map $\Omega \rightarrow C_0 : \omega \mapsto B(\omega)$, so $\mathcal{F}_t = \sigma(B)$. Let $F(\sigma) := \mathbb{E}\{f|B = \sigma\}$ for almost all $\sigma \in C_0$, so $f = F \circ B$. $F \in L^2$ since $(B)_*(\mathbb{P}) = \gamma$, Wiener measure. Take α^F so that

$$F(\omega) = \int_{C_0} F d\gamma + \int_0^T \alpha^F(\sigma)_s d\sigma(s).$$

Therefore,

$$f(\omega) = F(B(\omega)) = \mathbb{E}f + \int_0^T \alpha^F(\sigma)_s d\sigma(s) \Big|_{\sigma=B(\omega)}.$$

But the composition

$$\Omega \xrightarrow{B(-)} C_0 \xrightarrow{\int_0^T \alpha^F d\sigma} \mathbb{R}$$

is

$$\omega \mapsto \left(\int_0^T \alpha^F(B(-))_s dB_s \right) (\omega)$$

since it holds is α^F is elementary. Therefore, take $a^f(\omega)_s = \alpha^F(B(\omega))_s$. Uniqueness follows as before. \square

Corollary 9.16. (Martingale Representation of Brownian Motion.) *For B a BM on \mathbb{R} , suppose that M is an \mathcal{F}_*^B -martingale with $M_T \in L^2$. Then there is a unique $\alpha \in L^2(B)$ such that*

$$M_t = M_0 + \int_0^t \alpha_s dB_s \text{ almost surely}$$

for $0 \leq t \leq T$.

Proof. Take $\alpha_s = a_s^{M_T}$ as in Theorem 9.15. Hence,

$$M_T = \mathbb{E}M_T + \int_0^T a_s^{M_T} dB_s.$$

Therefore, by the martingale property of Itô integrals,

$$M_t = \mathbb{E}\{M_T | \mathcal{F}_t^B\} = \mathbb{E}M_T + \int_0^t a_s^{M_T} dB_s.$$

Finally, note that $\mathcal{F}_0^B = \{\emptyset, \Omega\}$, since $B_0(\omega) = 0$ for all $\omega \in \Omega$. Therefore, $\mathbb{E}\{-|\mathcal{F}_0^B\} = \mathbb{E}$, and so

$$\mathbb{E}M_T = \mathbb{E}\{M_T|\mathcal{F}_0^B\} = M_0.$$

Uniqueness holds just for M_T by Theorem 9.15. \square

Remark 9.17. We claimed (but did not prove) that we can choose $\left(\int_0^t \alpha_s dB_s\right)(\omega)$ for each t to make it continuous in t for all $\omega \in \Omega$. Therefore, if $\{M_t\}_{t \in [0, T]}$ is as above, there exists $\{M'_t\}_{t \in [0, T]}$ so that

- it is continuous in t for all $\omega \in \Omega$;
- $M'_t = M_t$ almost surely for each $t \in [0, T]$ (they disagree on sets of measure zero in \mathcal{F}_T);

so M'_t is $\mathcal{F}_t^B \vee Z_T$ measurable for each t , where Z_T is the collection of sets of measure zero in \mathcal{F}_T . In fact, $\{M'_t\}_{t \in [0, T]}$ will be a $\mathcal{F}'_t := \mathcal{F}_t^B \vee Z_T$ -martingale. (See [RW1] and [RW2].)

Example 9.18. Consider $\Omega = [0, 1]$ with Lebesgue measure λ^1 and the process $a : [0, 1] \times \Omega \rightarrow \mathbb{R}$ given by

$$a_t(\omega) := \begin{cases} t & t \neq \omega \\ 0 & t = \omega \end{cases}$$

Set $a'_t(\omega) := t$ for all $\omega \in \Omega$. $a'_t(\omega) = a_t(\omega)$ almost surely for each $t \in [0, 1]$, but it is not true that $a_t = a'_t$ for all $t \in [0, 1]$ almost surely. Note, however, that $a = a'$ in $L^2([0, 1] \times \Omega; \mathbb{R})$.

Corollary 9.19. *If M is an \mathcal{F}_*^B -martingale with M_T in L^2 , there exists a process $\langle M \rangle : [0, T] \times \Omega \rightarrow \mathbb{R}_{\geq 0}$ such that*

- (i) *it is adapted;*
- (ii) $\langle M \rangle_0 = 0$;
- (iii) $t \mapsto \langle M \rangle_t(\omega)$ *is non-decreasing;*
- (iv) $\{(M_t)^2 - \langle M \rangle_t | 0 \leq t \leq T\}$ *is an \mathcal{F}_*^B -martingale.*

Definition 9.20. $\langle M \rangle$ is the *increasing process* of M , or its *quadratic variation*.

Proof. (For M bounded, i.e., $\exists C$ such that $|M_t(\omega)| \leq C$ for all t and ω . This is not the case for BM.) For some $\alpha \in L^2(B)$,

$$M_t = M_0 + \int_0^t \alpha_s dB_s.$$

Apply Itô's formula

$$\text{“}M_t^2 = M_0^2 + \int_0^t 2M_t dM_t + \frac{1}{2} \int_0^t 2 dM_t dM_t\text{”}$$

with $\theta : \mathbb{R} \rightarrow \mathbb{R}$ given by $\theta(x) = x^2$, so $M_t^2 = \theta(M_t)$:

$$\text{“}\theta(x_t) = \theta(x_0) + \int_0^t \theta'(x_s) dx_s + \frac{1}{2} \int_0^t \theta''(x_s) dx_s dx_s\text{”}$$

So

$$M_t^2 = M_0^2 + 2 \int_0^t M_s \alpha_s dB_s + \frac{1}{2} \int_0^t \alpha_s^2 ds.$$

Hence,

$$M_t^2 - \int_0^t \alpha_s^2 ds = M_0^2 + 2 \int_0^t M_s \alpha_s dB_s,$$

and since M is bounded, $M \cdot \alpha \in L^2(B)$, and so the RHS is an \mathcal{F}^B -martingale. (M_0 is constant since it is \mathcal{F}_0^B -measurable, and $\mathcal{F}_0^B = \{\emptyset, \Omega\}$.) Now set $\langle M \rangle_t := \int_0^t \alpha_s^2 ds$. \square

Example 9.21. If $M_t = B_t$ then $B_t^2 - t = 2 \int_0^t B_s dB_s$, so $\langle B \rangle_t = t$ almost surely. This actually characterizes BM.

9.2 Chaos Expansions

Suppose that $f : C_0 \rightarrow \mathbb{R}$ is in L^2 . For some $\alpha \in L^2(B)$,

$$f(\sigma) = \int_{C_0} f + \int_0^T \alpha_t(\sigma) d\sigma(t) \text{ almost surely,}$$

with $B_t(\sigma) = \sigma(t)$. $\alpha : [0, T] \times \Omega \rightarrow \mathbb{R}$ has $\alpha_t \in L^2(C_0; \mathbb{R})$ for almost all $t \in [0, T]$ and is \mathcal{F}_t -measurable, therefore, there exists $\{\alpha_{s,t} | 0 \leq s \leq t\}$ in $L^2(B|_{[0,t]})$ such that

$$\alpha_t = \int_{C_0} \alpha_t + \int_0^t \alpha_{s,t}(\sigma) d\sigma(s) \text{ almost surely}$$

Therefore, writing $\bar{\alpha}_t := \int_{C_0} \alpha_t$,

$$f(\sigma) = \int_{C_0} f + \int_0^T \bar{\alpha}_t d\sigma(t) + \int_0^T \int_0^t \alpha_{s,t}(\sigma) d\sigma(s) d\sigma(t).$$

$\alpha_{s,t}$ is \mathcal{F}_s -measurable and in L^2 , so we repeat the above. Thus

$$\begin{aligned} f(\sigma) &= \bar{f} + \int_0^T \bar{\alpha}_{t_1}^{(1)} d\sigma(t_1) \\ &\quad + \int_0^T \int_0^{t_2} \bar{\alpha}_{t_1, t_2}^{(2)} d\sigma(t_1) d\sigma(t_2) \\ &\quad + \dots \\ &\quad + \int_0^T \int_0^{t_k} \dots \int_0^{t_2} \alpha_{t_1, \dots, t_k}^{(k)}(\sigma) d\sigma(t_1) \dots d\sigma(t_k) \text{ a.s.,} \end{aligned}$$

where $\bar{\alpha}^{(k-1)} \in L^2(\{0 \leq t_1 \leq \dots \leq t_{k-1} \leq T\} \subseteq [0, T]^{k-1}; \mathbb{R})$ and $\alpha_{t_1, \dots, t_k}^{(k)}$ is \mathcal{F}_{t_1} -measurable in $L^2(\{0 \leq t_1 \leq \dots \leq t_k \leq T\} \times \Omega; \mathbb{R})$. This corresponds to an orthogonal decomposition of $L^2(C_0; \mathbb{R})$, the *Wiener homogeneous chaos decomposition*. (All of the terms with an $\bar{\alpha}$ are orthogonal to the others.)

Example 9.22. $\mathbb{E} \int_0^T \int_0^t a_{s,t} d\sigma(s) d\sigma(t) \int_0^T b_s d\sigma(s) = 0$. By the isometry property,

$$\begin{aligned} \text{LHS} &= \int_0^T \mathbb{E} \left(\int_0^t a_{s,t} d\sigma(s) b_t \right) dt \\ &= \int_0^T b_t \mathbb{E} \left(\int_0^t a_{s,t} d\sigma(s) \right) dt \\ &= 0 \end{aligned}$$

since the expectation of an Itô integral is zero. This leads to the notion of Fock spaces in quantum field theory.