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 November 12, 2006

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

Contents

References

[BP] Czesław Bessaga & Alexsander 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: Itō 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 $\int_{\text{maps } \mathbb{R}^p \to \mathbb{R}^q}$.
- 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 $\{\Omega, \mathscr{F}\}$ where Ω is a set and $\mathscr{F}$ is a *σ-algebra* on Ω, so $\emptyset \in \mathscr{F}$, $A \in \mathscr{F} \implies \Omega \setminus A \in \mathscr{F}$ and $A_1, A_2, \dots \in \mathscr{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathscr{F}$.

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 $\{X, \mathscr{A}\}$ and $\{Y, \mathscr{B}\}$ are measurable spaces, $f : X \to Y$ is *measurable* if $B \in \mathscr{B} \implies f^{-1}(B) \in \mathscr{A}$. In general, $\sigma(f) := \{f^{-1}(B) | B \in \mathscr{B}\}$ 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 $\{Y, \mathscr{B}\}$.

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

Definitions 1.4. A *measure space* is a triple $\{\Omega, \mathscr{F}, \mu\}$ with $\{\Omega, \mathscr{F}\}$ a measure space and μ a *measure* on it, i.e. a $\mu : \mathscr{F} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ such that $\mu(\emptyset) = 0$ and $A_1, A_2, \dots \in \mathscr{F}$ 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 $\mu(\Omega) < \infty$. $\{\Omega, \mathscr{F}, \mu\}$ is a *probability space* (or μ is a *probability measure*) if $\mu(\Omega) = 1$.

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

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

(ii) Show that if $f, g : \Omega \to X$ 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 $\lambda^n(\text{rectangle}) =$ 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 $T_v : \mathbb{R}^n \to \mathbb{R}^n$ 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 $\{X, \mathscr{A}\}$: $c(A) := \#A$. Counting measure on $\mathbb{R}^n$ is also translation-invariant.

(iii) Dirac measure on any $\{X, \mathscr{A}\}$: 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 $\mathbb{L}(E;F) := \{\text{continuous linear maps } E \to F\}$ 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 $\{\Omega, \mathscr{F}\}$ 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 $\mathbb{R}^n$:

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

for $A \in \mathcal{B}(\mathbb{R}^n)$. Here $\mathrm{d}x = \mathrm{d}\lambda^n(x)$ and $\|x\|^2 = x_1^2 + \cdots + x_n^2$ 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 $\{\Omega, \mathscr{F}, \mu\}$, let $f : \Omega \to \Omega$ 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 $A \in \mathscr{F}$.

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 infinte-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* $\hat{\mu} : E^* \to \mathbb{C}$ is given by

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

for $\ell \in E^*$, where $i = \sqrt{-1} \in \mathbb{C}$. It exists since $\int_E |e^{i\ell(x)}|\,\mathrm{d}\mu(x) = \int_E \mathrm{d}\mu(x) = \mu(E) = 1 < \infty$. In fact, for all $\ell \in E^*$, $|\hat{\mu}(\ell)| \le 1$ 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 $H^* \to H : \ell \mapsto \ell^\sharp$, where $\ell^\sharp \in H$ has $\langle \ell^\sharp, x\rangle = \ell(x)$ for all $x \in H$. Therefore, we can consider $\hat{\mu}^\sharp : H \to \mathbb{C}$ given by

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

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

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

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

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 $x_0 \in E$, $\hat{\mu} : E^* \to \mathbb{C}$ is $\hat{\mu}(\ell) = \int_E e^{i\ell(x)}\,\mathrm{d}\delta_{x_0}(x) = e^{i\ell(x_0)}$. 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 $\{X, \mathscr{A}, \mu\}$, $\{Y, \mathscr{B}\}$ and $\theta : X \to Y$ measurable, giving $\theta_*\mu$ on Y, let $f : Y \to \mathbb{R}$ be measurable. Then $\int_X f\circ\theta\,\mathrm{d}\mu = \int_Y f\,\mathrm{d}(\theta_*\mu)$, in the sense that if one exists then so does the other and there is equality.*

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

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

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

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

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

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

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

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 $\ell \in F^*$,

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

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 $f : E \to \mathbb{C}$ 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 $f : V \to \mathbb{C}$ 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 $A^\intercal = \bar{A}$ 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, $\hat{\mu} : E^* \to \mathbb{C}$ is of positive type with $\hat{\mu}(0) = 1$ and is continuous on E^*.*

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

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

and is clearly Hermitian since

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

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

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 $k : F^* \to \mathbb{C}$ 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 $\mathbb{R}^n$:

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

This is a probability measure, since

$$\begin{aligned}\int_{\mathbb{R}^n} e^{-\|x\|^2/2} \,\mathrm{d}x &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-(x_1^2+\cdots+x_n^2)/2} \,\mathrm{d}x_1 \ldots \mathrm{d}x_n \\ &= \prod_{j=1}^{n} \int_{-\infty}^{\infty} e^{-x_j^2/2} \,\mathrm{d}x_j\end{aligned}$$

Also

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

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

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

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

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

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 $e^{-\frac{1}{2}\langle Cx,x\rangle} \leq e^{-\frac{1}{2}\lambda\|x\|^2} \,\mathrm{d}x$, and so all the above integrals exist.

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

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

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

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

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

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

since $\frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0}\det(I+hK) = \operatorname{tr} K$ 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 $\langle Ax,x\rangle = x_ix_j + x_jx_i$. □

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 $u:\mathbb{R}^n\to V$ with $\langle u(x),u(y)\rangle^\sim = \langle x,y\rangle$, so $u(x)=x_1e_1+\cdots+x_ne_n$ 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 $u:\mathbb{R}^n\to V$ by $u(x):=(\sqrt{C})^{-1}x$, so $\langle u(x),u(y)\rangle^\sim = \langle x,y\rangle$. By definition, $\lambda^{\langle\cdot,\cdot\rangle^\sim} := u_*\lambda^n$.

$$\int_V f\,\mathrm{d}\lambda^{\langle\cdot,\cdot\rangle^\sim} = \int_{\mathbb{R}^n} f(u(x))\,\mathrm{d}\lambda^n(x)$$
$$= \int_{\mathbb{R}^n} f\left(\sqrt{C}^{-1}x\right)\mathrm{d}\lambda^n(x)$$
$$= \det\sqrt{C}\int_{\mathbb{R}^n} f(y)\,\mathrm{d}y.$$

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

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

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

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

Remarks 3.5. (i) If $\dim V=n$ and $u:\mathbb{R}^n\to V$ 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

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

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 $T:H\to V$ 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 $\tilde{T}:=T|_{(\ker T)^\perp}:(\ker T)^\perp\to V$ 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,*

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

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

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

For $\alpha \in C$, 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 $\mathbb{C}$. □

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

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

Recall that if $\{V, \langle\cdot,\cdot\rangle_V\}$ is a real Hilbert space with $\dim V \le \infty$, then V^* has a natural inner product making it a Hilbert space, with Riesz isometric isomorphism $V^* \to V : \ell \mapsto \ell^\sharp$ 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. □

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 $T : \mathbb{R}^n \to V$, then, for $\ell \in V^*$,

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

by Remark 2.6. Moreover,

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

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.

($\Leftarrow$) If $\hat{\mu}(\ell) = e^{-\frac{1}{2}(\|\ell\|')^2}$ for some $\langle\cdot,\cdot\rangle'$ on V^*, take $T : \mathbb{R}^m \to \{V, \langle\cdot,\cdot\rangle^\sim\}$ 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 $\ell \in V^*$,

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

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

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 $\hat{\mu}(\ell) = e^{-\frac{1}{2}\|\ell\|^2_{L^2}}$.*

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

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

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

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

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

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 $\hat{\gamma}(\ell) = e^{-\frac{1}{2}\|\ell\|_{L^2}^2}$.*

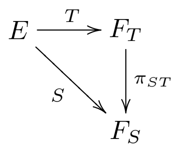
The proof of this mimics that of Theorem 3.13 and is omitted. However, we have not yet established whether or not there are any non-degenerate Gaussian measures on infinite-dimensional spaces!

5 Cylinder Set Measures

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

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

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



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 $T \in \mathscr{A}(E)$. 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 $\{\mu_T\}_{T\in\mathscr{A}(E)}$ 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 $\text{id} : E \to E$ the identity map.

(iii) A real Hilbert space $\{H, \langle\cdot,\cdot\rangle_H\}$. We have a canonical Gaussian CSM on H, $\{\gamma_T^H | T \in \mathscr{A}(H)\}$. γ_T^H on F_T (where $T : H \to F_T$ 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

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

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

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

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

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

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 $\ell^* : \mathbb{R} \to H \cong H^*$ 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

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

$\tilde{\ell}^{-1} : \mathbb{R} \to (\ker \ell)^\perp \subseteq H$ 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. $\square$

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

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

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

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

(i) $\widehat{\gamma_\cdot^H}(\ell) = \exp(-\frac{1}{2}\|\ell^\sharp\|_H^2)$*;*

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

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

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

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

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

where $\widetilde{S \circ \theta} : E_1 \to \tilde{F}$ 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 $\theta_*(\mu_\cdot)$ 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 $\{\theta_*(\gamma_\cdot^H)_T\}_T$ is a measure on G, i.e. θ radonifies the canonical Gaussian CSM $\{\gamma_T^H\}_{T\in\mathscr{A}(H)}$ on H.

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

(ii) If $\mathrm{id} : E \to E$ 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 $i : H \to E$ is a continuous linear injective map with dense range that γ-radonifies we say that $i : H \to E$ is an *abstract Wiener space* (or AWS). For example, $L^2 \to L^1$. The measure induced on E is called the *abstract Wiener measure* of $i : H \to E$.

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 $\ell \in E^*$:

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

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

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

Example 5.12. Classical Wiener space. Let

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

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

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

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

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

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

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

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

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 $\{\mu_T\}_{T \in \mathscr{A}(E)}$ is a* CSM *on E, then*

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

for all $\ell \in G^*$. In particular, if $T \in \mathscr{A}(E)$, then $\widehat{\mu_T}(\ell) = \hat{\mu}(T^*(\ell))$ 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 $= 1$ (probabilty measure) but LHS $= 1$ since $\theta_*(\mu.) = \delta_0$. □

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

Proof. Given $f : E^* \to \mathbb{C}$ of positive type and $T \in \mathscr{A}(E)$ (so that $T : E \to F_T$ is surjective), the composition $f \circ T^* : F_T^* \to \mathbb{C}$ 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 $\{\mu_T\}_{T \in \mathscr{A}(E)}$ forms a CSM.

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

Given a CSM $\{\mu_T\}_{T \in \mathscr{A}(E)}$ 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 $F = \operatorname{span}\{\ell_1, \dots, \ell_n\}$, then $\widehat{\mu.}$ is continuous on F.

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

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

(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. □

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 $T \in \mathscr{A}(E)$ then $\mu = \nu$;

(ii) if $\hat{\mu} = \hat{\nu}$ 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* $\operatorname{Cyl}(E) := \{T^{-1}(B) | B \in \mathcal{B}(F_T), T \in \mathscr{A}(E)\}$. This is an algebra of subsets, but not a σ-algebra if $\dim E = \infty$. Given a probability measure μ on E and $T \in \mathscr{A}(E)$, $A = T^{-1}(B)$ for some $B \in \mathcal{B}(F_T)$, $\mu(A) = T_*(\mu)(B) = \mu_T(B)$. Therefore, the CSM $\{T_*\mu | T \in \mathscr{A}(E)\}$ determines $\mu(A)$ for all $A \in \operatorname{Cyl}(E)$. Thus, the theorem follows from the following Lemma 5.18: □

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

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

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

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;
- 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 $\overline{B_\varepsilon(\sigma_0)} \in \sigma(\operatorname{Cyl}(E))$. 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

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

because

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

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

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

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

6.1 Construction of the Paley-Wiener Integral

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

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

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

Proof. (i)

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

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. □

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 $\|\alpha(x)\|_G \le k\|x\|_F$ 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, $\|\tilde{\alpha}\| \le k$. 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 $x_n \to x$ in F. Then

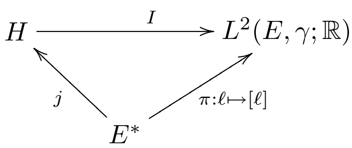
$$\|\alpha(x_n) - \alpha(x_m)\|_G = \|\alpha(x_n - x_m)\|_G \le k\|x_n - x_m\|_F$$

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

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

Therefore, $\|\tilde{\alpha}\| \le k$ 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 $x \in F$. □

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 $I : H \to L^2(E, \gamma; \mathbb{R})$, with $I(h) := \langle h, -\rangle_H^\sim$, such that*



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

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

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

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

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

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

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

If $h \in H =$ the image of j, take ℓ_n such that $\ell_n^\sharp = h$ 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 $\langle h, -\rangle_H^\sim$ 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 $I(h) = \langle h, -\rangle_H^\sim$ — 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 $H = L_0^{2,1}([0, T]; \mathbb{R}^n)$,

$$\begin{aligned}
\langle h^1, h^2\rangle_H &= \int_0^T \langle \dot{h}^1(s), \dot{h}^2(s)\rangle_{\mathbb{R}^n} \, \mathrm{d}s \\
&= \text{“} \int_0^T \langle \dot{h}^1(s), \mathrm{d}h^2(s)\rangle_{\mathbb{R}^n} \text{” (Stieltjes)}
\end{aligned}$$

"$\mathrm{d}h(s)$" means "$\dot{h}(s)\,\mathrm{d}s$".

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

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

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

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

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.

$$H = L_0^{2,1} \underset{\int_0^\cdot - \mathrm{d}s}{\overset{\frac{\mathrm{d}}{\mathrm{d}t}}{\rightleftarrows}} L^2$$

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 $i : H \to E$, *if* $h \in H$, $h \neq 0$, *then*

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

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 $\ell_*(\gamma) = (\ell \circ i)_*(\gamma^H) = \gamma^H_{\langle h, - \rangle_H}$, as required.

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

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

But $\widehat{(\ell_n)_*\gamma}(s) = \hat{\gamma}(s\ell_n) = \int_E e^{is\ell_n(x)} \, \mathrm{d}\gamma(x)$. Now $\ell_n \to I(h)$ 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 $e^{ix} - e^{iy} = 2ie^{i(x+y)/2} \sin \frac{x-y}{2}$

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

So

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

and the result follows from Bochner's Theorem. □

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

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

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

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

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

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

(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 $L^2([0, T]; \mathbb{R}) \to L^2(C_0, \gamma; \mathbb{R})$ given by

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

an isometry onto its image. □

6.2 The Structure of Gaussian Measures

Definition 6.11. If $\{\Omega, \mathscr{F}, \mathbb{P}\}$ is a probability space, G a separable Banach space, and $f : \Omega \to G$, 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. $I(h) : E \to \mathbb{R}$ is a Gaussian random variable on $\{E, \mathcal{B}(E), \gamma\}$ if $i : H \to E$ is an AWS and $h \in H$, by Proposition 6.9.

Remark 6.13. Let $\{\Omega, \mathscr{F}, \mu\}$ be a measure space, $\{X, d\}$ a metric space, and $f_j, g : \Omega \to X$ measurable functions for $j \in \mathbb{N}$. $f_j \to g$ *almost everywhere/almost surely/with probability 1* means that there is a set $Z \in \mathscr{F}$ with $\mu(Z) = 0$ such that $f_j(x) \to g(x)$ as $j \to \infty$ 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 $\{\Omega, \mathscr{F}, \mathbb{P}\}$ be a probability space and $f_j : \Omega \to \mathbb{R}$ a sequence of Gaussian random variables such that $f_j \to 0$ almost surely as $j \to \infty$. Then $f_j \to 0$ in L^2. In particular, every Gaussian $\mathbb{R}$-valued random variable lies in $L^2(\Omega, \mathscr{F}, \mathbb{P}; \mathbb{R})$.*

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

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

So $\widehat{\gamma_j}(s) \to 1$ for all $s \in \mathbb{R}$ (✠). Now,

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

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, $c_j^{-1} \to 0$ as $j \to \infty$ by (✠). But

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

□

Remark 6.15. From the proof we saw that if $f : \Omega \to \mathbb{R}$ 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 $i : H \to E$ such that $i : H \to E$ is an* AWS *with $\gamma = i_*(\gamma^H_\cdot)$.*

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 $j : E^* \to L^2(E, \gamma; \mathbb{R})$ 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 $E^* \to H$. By definition, this is linear with dense range. So see that it is continuous, let $\ell_n \to \ell$ in E^* as $n \to \infty$. Then

(i) $\ell_n - \ell \to 0$ in E^* and $\ell_n - \ell$ is a Gaussian random variable on $\{E, \gamma\}$;

(ii) $\ell_n(x) - \ell(x) \to 0$ for all $x \in E$, so $\ell_n - \ell \to 0$ (almost) surely.

Therefore, by Lemma 6.14, $\ell_n - \ell \to 0$ 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 $i := j^* : H \cong H^* \to E^{**}$ by $i(h)(\ell) := \langle h, j(\ell)\rangle_H$ for $h \in H$, $\ell \in E^*$ (†).

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

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

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

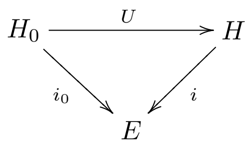
$$\begin{aligned}
\widehat{i_*\gamma^H}(\ell) &= \widehat{\gamma^H}(i^*(\ell)) \\
&= \widehat{\gamma^H}(j(\ell)) \\
&= e^{-\frac{1}{2}\|j(\ell)\|_H^2} \text{ by Proposition 5.6} \\
&= e^{-\frac{1}{2}\|\ell\|_{L^2}^2} \text{ by the definition of } j
\end{aligned}$$

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

(ii) i is injective. This is easy, since j has dense range.

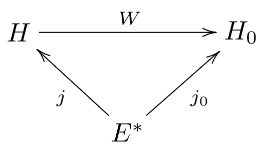
(iii) i has dense range. It is enough to show that j is injective. Suppose that $\ell \in \ker j$, so $j(\ell) = 0$. Then $\ell = 0$ almost surely. But if $\ell \neq 0 \in E^*$ then $\ker \ell$ is a proper closed subspace of E, so there exist $x \in E$ and $r > 0$ with $B_r(x) \cap \ker \ell = \emptyset$.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

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



Proof. Take $j : E^* \to H$ and $j_0 : E^* \to H_0$ 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 $W_0 : j(E^*) \to j_0(E^*)$ 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 $W : H \to H_0$ 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 $W^*W = WW^* = \mathrm{id}$. Also,



So take $U = W^* : H_0 \to H$, 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 $i_0 : H_0 := L^2([0,T];\mathbb{R}) \to C_0$, where $i_0(h)(t) = \int_0^t h(s)\,\mathrm{d}s$ for $0 \le t \le T$, $h \in H_0$. We have $U : L^2 \to L_0^{2,1}$ as $U(h)(t) = \int_0^t h(s)\,\mathrm{d}s$ for $0 \le t \le T$, $h \in L^2$.

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

Let $i : H \to E$ be an AWS with measure γ. Consider $T_h : E \to E$ 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 $i = \mathrm{id}$. Recall that

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

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

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

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

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

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

Recall that if μ, ν on $\{X, \mathscr{A}\}$ 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 $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} : X \to \mathbb{R}_{\geq 0}$ such that $\nu = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mu$, i.e.

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

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

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

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

Proof.

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

□

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

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

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

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

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

$$\begin{aligned}\int_E (e^{\langle h,-\rangle^{\sim}_H})^p \,\mathrm{d}\gamma &= \int_{\mathbb{R}} t^p \,\mathrm{d}(I(h)_*\gamma)(t) \\ &= \int_{\mathbb{R}} t^p \,\mathrm{d}\gamma^H_{\langle h,-\rangle_H}(t) \\ &< \infty \text{ for } 1 \leq p < \infty\end{aligned}$$

since $\mathrm{d}\gamma^H_{\langle h,-\rangle_H}(t) = \frac{1}{N}e^{-ct^2}\,\mathrm{d}t$ 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 $\int_E e^{\rho\langle g,-\rangle^{\sim}_H(x)+z\langle h,-\rangle^{\sim}_H(x)}\,\mathrm{d}\gamma(x) = \int_E e^{i\langle ag+bh,-\rangle^{\sim}_H}\,\mathrm{d}\gamma$ and $ag+bh \in H$ since H is a real Hilbert space. So $\int_E e^{\rho\langle g,-\rangle^{\sim}_H(x)+z\langle h,-\rangle^{\sim}_H(}$ $e^{-\frac{1}{2}\|ag+bh\|^2_H}$ from before, as required.

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

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

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

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

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

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

This is at most $\mathrm{length}(\sigma)\int_\Omega \sup_{z\in\mathscr{C}}|F(z,\omega)|\,\mathrm{d}\mu(\omega)$, where $\mathrm{length}(\sigma) := \int_0^T |\dot\sigma(t)|\,\mathrm{d}t$. 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,

$$\begin{aligned}|F(z,\omega)| &= e^{\mathrm{Re}\,\rho\langle g,-\rangle^{\sim}_H(\omega)+\mathrm{Re}\,z\langle h,-\rangle^{\sim}_H(\omega)} \\ &= e^{\langle(\mathrm{Re}\,\rho)g+(\mathrm{Re}\,z)h,-\rangle^{\sim}_H(\omega)}\end{aligned}$$

for $\omega \in E$, and

$$\int_0^T\int_E |\dot\sigma(t)||F(\sigma(t),\omega)|\,\mathrm{d}\gamma(\sigma)\mathrm{d}t = \int_0^T\int_E e^{\langle k(t),-\rangle^{\sim}_H(\omega)}\,\mathrm{d}\gamma(\omega)|\dot\sigma(t)|\,\mathrm{d}t < \infty,$$

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

$$\int_E e^{\langle k(t),-\rangle^{\sim}_H}\,\mathrm{d}\gamma = \int_{\mathbb{R}} e^s\,\mathrm{d}\gamma^H_{\langle k(t),-\rangle_H}(s) < \infty.$$

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

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

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 $\ell \in E^*$,

$$\begin{aligned}\widehat{\gamma_h} &= e^{\sqrt{-1}\ell(i(h))}\hat\gamma(\ell) \\ &= e^{\sqrt{-1}\ell(i(h))}e^{-\frac{1}{2}\|j(\ell)\|^2_H} \\ &= e^{\sqrt{-1}\langle j(\ell),h\rangle^{\sim}_H-\frac{1}{2}\|j(\ell)\|^2_H}.\end{aligned}$$

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

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

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

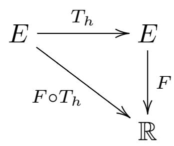
Theorem 7.7. (Integrated Cameron-Martin.) *If $F : E \to \mathbb{R}$ (or $E \to$ any separable Banach space) is measurable and $h \in H$ then*

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

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

Proof. By Theorem 7.5 and Proposition 2.4,

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



$\square$

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

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

Formally differentiate at $t = 0$:

$$\int_E \mathrm{D}F(x)(i(h)) \,\mathrm{d}\gamma(x) = \int_E F(x)\langle h,-\rangle^{\sim}(x) \,\mathrm{d}\gamma(x).$$

If F is a "nice" differentiable function with derivative $\mathrm{D}F : E \to \mathbb{L}(E;\mathbb{R})$, we have the above integration by parts formula.

(ii) In $\mathbb{R}$,

$$\int_{\mathbb{R}} f'(x)v(x) \,\mathrm{d}x = -\int_{\mathbb{R}} f(x)v'(x) \,\mathrm{d}x$$

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

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

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

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

We want to allow more general vector fields $V : E \to E$. 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 $x \mapsto x + v \iff v \in i(H)$.

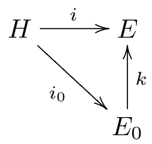
Theorem 7.10. *If H is an infinite-dimensional separable Hilbert space then the canonical Gaussian* CSM *on H, $\{\gamma_\cdot^H\}$, is not a measure on H.*

Proof. Suppose not, so that $\gamma_\cdot^H = \gamma$ is a measure on H. Then $i = \mathrm{id} : H \to H$ 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. □

Remarks 7.11. For $i : H \to E$ 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 $\exists i_0 : H \to E_0$, also an AWS, and $k \in \mathbb{L}(E_0; E)$ injective and compact such that $k \circ i_0 = i$:



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 $\{X, \mathscr{A}\}$ is a map $z : S \times \Omega \to X$ for some probability space $\{\Omega, \mathscr{F}, \mathbb{P}\}$ such that for all $s \in S$, the map $\Omega \to X : \omega \mapsto z_s(\omega) := z(s, \omega)$ is measurable.

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

In the Kolmogorov model of probability theory, the probability that our system or process behaves is 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 $A_i \in \mathscr{A}$ for $i = 1, \ldots, k$ and $s_1, \ldots, s_k \in S$ then the probability that the process has value in A_i at $s = s_i$ for each i is $\mathbb{P}\{\omega \in \Omega | z_{s_i}(\omega) \in A_i \text{ for } i = 1, \ldots, k\}$, which we often write as $\mathbb{P}\{z_{s_i} \in A_i \text{ for } i = 1, \ldots, k\}$ or $\mathbb{P}_{\boldsymbol{s}}(A_1 \times \cdots \times A_k)$, where $\mathbb{P}_{\boldsymbol{s}}$ is the push-forward measure $(z_{\boldsymbol{s}})_*(\mathbb{P})$ on X^k, where

$$\begin{aligned} z_{\boldsymbol{s}} : \Omega &\to X^k \\ \omega &\mapsto (z_{s_1}(\omega), \ldots, z_{s_k}(\omega)), \end{aligned}$$

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

Example 8.3. If $i : H \to E$ is an AWS with measure γ, consider $z : H \times E \to \mathbb{R}$, $\{E, \mathcal{B}(E), \gamma\}$ our measure space, given by $z(h, \omega) = \langle h, - \rangle_H^\sim(\omega)$ 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 $\mathscr{A} = \mathcal{B}(X)$) we say that a stochastic process z is *continuous* (or *sample continuous*) if the map $S \to X : s \mapsto z_s(\omega)$ is continuous for all $\omega \in \Omega$. (Some authors allow almost all $\omega \in \Omega$.)

Example 8.5. Let $\Omega = C_0 = C_0([0, T]; \mathbb{R}^n) = \{\text{continuous paths in } \mathbb{R}^n \text{ starting at } 0\}$, $\mathbb{P} =$ Wiener measure, $X = \mathbb{R}^n$, $\mathscr{A} = \mathcal{B}(\mathbb{R}^n)$, $S = [0, T]$. Any measurable vector field $V : C_0 \to C_0$ determines a sample continuous process $z : [0, T] \times C_0 \to \mathbb{R}^n$ 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 $z : [0, T] \times \Omega \to X$ is a continuous process, with X a separable Banach space, we get $\Phi : \Omega \to C([0, T]; X)$ 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* $\mathscr{L}_z$ of z is the push-forward measure $\mathscr{L}_z := \Phi_*(\mathbb{P})$ 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 $[0, T] \times C([0, T]; X) \to X$ using the measure $\mathscr{L}_z$ on $C([0, T]; X)$ has the same finite-dimensional distributions as $z : [0, T] \times \Omega \to X$. Consequently, the finite-dimensional distributions of z determine its law.

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

(i) $B_0(\omega) = 0$ for all $\omega \in \Omega$;

(ii) B is sample continuous;

(iii) $\mathscr{L}_B$ 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 $B : [0, \infty) \times \Omega \to \mathbb{R}^n$ 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 $h \in H := L_0^{2,1}([0,T];\mathbb{R}) \subset C_0$ and B is a BM on $\mathbb{R}^n$, write $\int_0^T \langle \dot{h}(s), \mathrm{d}B_s \rangle_{\mathbb{R}^n} : \Omega \to \mathbb{R}$ for the composition

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

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

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

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

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

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

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

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

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

8.2 Construction of Itō's Integral

We want to construct the Itō integral

$$\int_0^T \langle a_s, \mathrm{d}B_s \rangle : \Omega \to \mathbb{R},$$

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

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

Note that in the Paley-Wiener integral $\int_0^t \langle f(s), \mathrm{d}B_s \rangle$ 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 $\{\Omega, \mathscr{F}, \mathbb{P}\}$ be a probability space. A *filtration* is a family of σ-algebras on Ω, $\{\mathscr{F}_t | t \in [0,T]\}$, such that

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

Example 8.16. Given a process $z : [0,T] \times \Omega \to X$, with $\{X, \mathscr{A}\}$ any measurable space, define $\mathscr{F}_t := \mathscr{F}_t^z = \sigma\{z_r : \Omega \to X | 0 \le r \le t\}$, the "events up to time t" or "the past at time t". $\mathscr{F}_*^z$ 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 $0 \le s_1 \le s_2 \le \cdots \le s_k \le t$ and $A_1, \ldots, A_k \in \mathcal{B}(\mathbb{R}^n)$,

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

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

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

Take $n = 1$ for ease of notation.

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

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

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

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

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

Let B be a BM on $\mathbb{R}$, $B : [0,T] \times \Omega \to \mathbb{R}$. Given $0 \le t_0 < t_1 < \cdots < t_k \le T$, 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

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

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

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

(i) if $i < j$, $\int_\Omega \alpha_i \alpha_j \Delta_i B \Delta_j B \,\mathrm{d}\mathbb{P} = 0$;

(ii) $\int_\Omega \alpha_i^2 (\Delta_i B)^2 \,\mathrm{d}\mathbb{P} = (\mathbb{E}\alpha_i^2)\Delta_i t$;

(iii) as an immediate consequence of (i) and (ii),

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

Assuming Proposition 8.20,

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

$$\left\| \int_0^T a_s \,\mathrm{d}B_s \right\|_{L^2} = \|a_\cdot\|_{L^2([0,T]\times\Omega;\mathbb{R})}.$$

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

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

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

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

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

i.e.,

$$\int_\Omega \left(\int_0^T a_s \, \mathrm{d}B_s(\omega) \right)^2 \mathrm{d}\mathbb{P}(\omega) = \int_0^T \mathbb{E}|a_s|^2 \, \mathrm{d}s.$$

Proof. Lemma 6.2. ☐

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

From this $\int_\Omega \bar{f} \,\mathrm{d}\mathbb{P} = \int_\Omega f \,\mathrm{d}\mathbb{P} < \infty$ 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 $\bar{f} = \overline{f^+} - \overline{f^-}$. 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 $\mathbb{E}\{|f||\mathscr{A}\}(\omega) = \overline{f^+}(\omega) + \overline{f^-}(\omega)$. Also, $|\mathbb{E}\{|f||\mathscr{A}\}(\omega)| = |\bar{f}(\omega)| = |\overline{f^+}(\omega) + \overline{f^-}(\omega)|$, giving (ii), since $\overline{f^+} \geq 0$, $\overline{f^-} \geq 0$ almost everywhere.

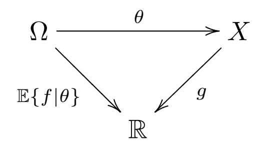
(iii) If $f \in L^1$,

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

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

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

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



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

- $g \circ \theta$ is $\sigma(\theta)$-measurable;
- if $A \in \mathscr{A}$ then

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

- 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 $\mathbb{E}\{f|\theta\} : \Omega \to \mathbb{R}$ 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 nth morning, i.e.

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

Consider just wind speed, for instance. Let $f_n : \Omega \to \mathbb{R}$ be the windspeed at mid-day on the nth day. Given some observation in the morning, we want to forecast f_n. We need $g_n : X \to \mathbb{R}$, which tells us that if on the nth 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 $L^2(\Omega, \mathscr{F}, \mathbb{P}; \mathbb{R})$ norm. Now, by Lemma 8.25 functions of the form $g_n \circ \theta_n$ in L^2 are exactly elements of $L^2\left(\Omega, \sigma(\theta_n), \mathbb{P}|_{\sigma(\theta_n)}; \mathbb{R}\right)$. 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. $\mathbb{E}\{f|\theta = x\} = \int_{\theta^{-1}(x)} f(y)\, \mathrm{d}\mathbb{P}_x(y)$, where $\mathbb{P}_x$ is a measure on the fibre $\theta^{-1}(x)$, known as the *disintegration of* $\mathbb{P}$.

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

(i) If $\mathscr{A}, \mathscr{B} \subseteq \mathscr{F}$, we say that $\mathscr{A}$ and $\mathscr{B}$ are *independent*, and write $\mathscr{A} \amalg \mathscr{B}$, if $A \in \mathscr{A}, B \in \mathscr{B} \implies \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

(ii) If for $j = 1, 2$, $\{X_j, \mathscr{A}_j\}$ are measurable spaces with $f_j : \Omega \to X_j$ 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 $f_1 \times f_2 : \Omega \to X_1 \times X_2 : \omega \mapsto (f_1(\omega), f_2(\omega))$ satisifies $(f_1 \times f_2)_*(\mathbb{P}) = (f_1)_*(\mathbb{P}) \otimes (f_2)_*(\mathbb{P})$, 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 $A_1 \in \mathscr{A}_1$ and $A_2 \in \mathscr{A}_2$, 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 $A_j \in \mathscr{A}_j$ 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 $F_j : X_j \to \mathbb{R}$ 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 $F_1 \circ f_1, F_2 \circ f_2 \in L^1(\Omega, \mathscr{F}, \mathbb{P}; \mathbb{R})$.

Proof. The left-hand side is

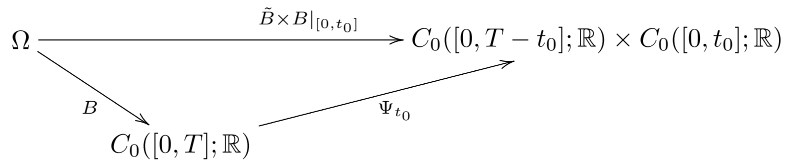
$$\int_{X_1 \times X_2} F_1(x)F_2(y)\,\mathrm{d}(f_1 \times f_2)_*(\mathbb{P})(x, y) = \int_{X_1} F_1(x)\,\mathrm{d}(f_1)_*(\mathbb{P})(x) \int_{X_2} F_2(y)\,\mathrm{d}(f_2)_*(\mathbb{P})(y)$$

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

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

Theorem 8.33. *Let $B : [0, T] \times \Omega \to \mathbb{R}$ 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., $\tilde{B} : \Omega \to C_0([0, T - t_0]; \mathbb{R})$ is independent of $B.|_{[0,t_0]} : \Omega \to C_0([0, t_0]; \mathbb{R})$. Also, both $\tilde{B}$ and $B.|_{[0,t_0]}$ are* BM*s on $\mathbb{R}$.*

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



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 $0 \le s \le t \le u \le v \le T$. Then if B is a* BM *on $\mathbb{R}$, $B_t - B_s$ is independent of $B_v - B_u$, i.e.,*

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

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

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

$$f \amalg \mathscr{A} \implies \mathbb{E}\{f | \mathscr{A}\} = \mathbb{E}f.$$

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

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

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

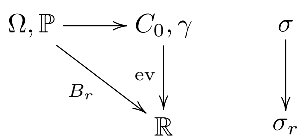
Theorem 8.36. (The Martingale Property of Brownian Motion.) *If $\mathscr{F}_s = \mathscr{F}_s^B = \sigma\{B_r | 0 \le r \le s\}$ for a* BM *B on $\mathbb{R}$, then for $0 \le s \le t$,*

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

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

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

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



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

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

□

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

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

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

□

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

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

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

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

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

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

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

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

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

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

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

(ii) Easy exercise. □

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

(i) $\mathbb{E}\alpha_i\alpha_j\Delta_iB\Delta_jB = 0$;

(ii) $\mathbb{E}\alpha_i^2(\Delta_iB)^2 = (\mathbb{E}\alpha_i^2)\Delta_it$.

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

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

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

(ii)

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

□

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

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

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

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

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

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

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

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

In particular, we have the *Itō isometry*:

$$\mathbb{E}\left(\int_0^T a_s \,\mathrm{d}B_s\right)^2 = \mathbb{E}\left(\int_0^T (a_s)^2 \,\mathrm{d}s\right).$$

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

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

is $\mathcal{B}[0,t] \circledast \mathscr{F}_t$-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 $\int_0^T a_s \,\mathrm{d}B_s$ and the Itō integral $\int_0^T a_s \,\mathrm{d}B_s$, defined in different ways. Later work will show that they agree.

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

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

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 $\int_0^t (a|_{[0,t]})_s \, \mathrm{d}(B|_{[0,t]})_s$, and since it clearly agrees with $\int_0^t a_s \, \mathrm{d}B_s$ for $a \in \mathscr{E}$, by continuity, it agrees for all $a \in L^2(B)$.

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

(i) it is adapted (i.e. $M_t : \Omega \to \mathbb{R}^n$ is $\mathscr{F}_t$-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 $s \le t$ then $\mathbb{E}\{M_t | \mathscr{F}_s\} = M_s$ almost surely.

If no filtration is specified then M is a *martingale* if it is an $\mathscr{F}_*^M$-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 $z : [0,T] \times \Omega \to \mathbb{R}$ defined by*

$$z_t(\omega) := \left(\int_0^t a_s \, \mathrm{d}B_s \right)(\omega)$$

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

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

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

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

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

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

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

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

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

Proof. Since $z_t := \int_0^t a_s \, \mathrm{d}B_s$ is an $\mathscr{F}_*^B$-martingale,

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

and $z_t \amalg \mathscr{F}_0^B$ since $\mathscr{F}_0^B = \{\emptyset, \Omega\}$. □

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

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

$$z_t(\omega) = a_t(\omega) + \left(\int_0^t \alpha_s \, \mathrm{d}B_s \right)(\omega),$$

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

$$\begin{aligned} \theta(z_t(\omega)) = \theta(z_0(\omega)) + \int_0^t \theta'(z_s(\omega)) a_s'(\omega) \, \mathrm{d}s + \left(\int_0^t \theta'(z_s(-)) \alpha_s \, \mathrm{d}B_s \right)(\omega) \\ + \frac{1}{2} \int_0^t \theta''(z_s(\omega)) \alpha_s(\omega) \alpha_s(\omega) \, \mathrm{d}s \text{ almost surely.} \end{aligned}$$

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 $0 = t_0 < t_1 < \cdots < t_{k+1} = t$. For "nice" θ,

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

The first and second terms in Itō'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

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

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

	$\mathrm{d}t$	$\mathrm{d}B$
$\mathrm{d}t$	0	0
$\mathrm{d}B$	0	$\mathrm{d}t$

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 Itō's formula. Let $\theta : \mathbb{R} \to \mathbb{R}$ be $\theta(x) = x^2$. So

$$\begin{aligned} B_t^2 &= B_0^2 + \int_0^t 2 B_s \, \mathrm{d}B_s + \frac{1}{2} \int_0^t 2 \, \mathrm{d}s \\ &= 0 + 2 \int_0^t B_s \, \mathrm{d}B_s + t. \end{aligned}$$

Thus,

$$\int_0^t B_s \, \mathrm{d}B_s = \frac{1}{2} B_t^2 - \frac{1}{2} t.$$

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

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

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

Claim 8.51. *M satisfies*

$$M_t = 1 + \int_0^t M_s \dot{h}(s)\,\mathrm{d}B_s$$

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

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

$$\begin{cases} dM_t = M_t \dot{h}_t\,\mathrm{d}B_t \\ M_0 = 1 \end{cases}$$

Proof. Set

$$z_t = \int_0^t \dot{h}(s)\,\mathrm{d}B_s - \frac{1}{2}\int_0^t |\dot{h}(s)|^2\,\mathrm{d}s$$

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

$$\begin{aligned}\theta(z_t(\omega)) = \theta(z.(\omega)) &+ \int_0^t e^{z_s(\omega)}\left(-\frac{1}{2}\dot{h}_s^2\right)\mathrm{d}s \\ &+ \left(\int_0^t e^{z_s(-)}\dot{h}_s\,\mathrm{d}B_s\right)(\omega) + \frac{1}{2}\int_0^t e^{z_s(\omega)}(\dot{h}_s)^2\,\mathrm{d}s\end{aligned}$$

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

$$M_t = 1 + \int_0^t M_s \dot{h}(s)\,\mathrm{d}B_s \text{ almost surely.}$$ ☐

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

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

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

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

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

Therefore,

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

But M_t is the composition

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

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

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

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

$$\begin{aligned}1 &= \int_{C_0} e^{p\langle h,-\rangle_H^{\sim} - \frac{1}{2}p^2\|h\|_H^2} \\ &= \mathbb{E}\{(M_t)^p\}e^{\frac{1}{2}p\|h\|_H - \frac{1}{2}p^2\|h\|_H^2}.\end{aligned}$$ ☐

Proposition 8.53. *M_t, $0 \le t \le T$, is an $\mathscr{F}^B_*$-martingale. In particular, $\mathbb{E}M_t = 1$ for all $0 \le t \le T$.*

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

$$\begin{cases} \mathrm{d}x_t = A(x_t)\,\mathrm{d}t + X(x_t)\,\mathrm{d}B_t \\ x_0 = q. \end{cases}$$

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

$$x_t = q + \int_0^t A(x_s)\,\mathrm{d}s + \int_0^t X(x_s)\,\mathrm{d}B_s \text{ almost surely.}$$

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

$$\frac{\mathrm{d}x_t}{\mathrm{d}t} = A(x_t).$$

9 Itō Integrals as Divergences

9.1 The Clark-Ocone Theorem and Integral Representation

We work with canonical 1-dimensional Brownian motion:

$$\begin{gathered}\Omega = C_0([0,T];\mathbb{R})\\ \mathscr{F}_t = \sigma\{\mathrm{ev}_s\,|0 \le s \le t\} = \sigma\{\rho \mapsto \rho(s)|0 \le s \le t\}\\ B_t(\omega) = \omega(t) = \mathrm{ev}_t(\omega)\end{gathered}$$

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

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

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

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

$$\int_{C_0} F(\sigma + V_\cdot(\sigma))e^{-\int_0^T \frac{\partial}{\partial t}V_t(\sigma)\,\mathrm{d}\sigma(t) + \frac{1}{2}\int_0^T |\frac{\partial}{\partial t}V_t(\sigma)|^2\,\mathrm{d}t}\,\mathrm{d}\gamma(\sigma) = \int_{C_0} F(\sigma)\,\mathrm{d}\gamma(\sigma),$$

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 $\mathscr{F}_0$-measurable and $\mathscr{F}_0 = \{\emptyset, \Omega\}$. Thus $V_t = t\alpha_0$ for $0 \le t \le T$, so $V_\cdot \in H := L_0^{2,1}([0,T];\mathbb{R})$ and we can apply the Cameron-Martin Formula, Theorem 7.7, with $\check{F} : C_0 \to \mathbb{R}$ given by $\check{F}(\sigma) := F(\sigma + V_\cdot)$. Thus

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

i.e.

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

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 $0 = t_0 < \cdots < t_n = T_0 < t_{n+1} = T$. We have

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

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

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

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

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

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,

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

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

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

Then

$$\int_{C_0} F(\sigma + V(\sigma)) e^{-\int_0^T \frac{\partial}{\partial s}V_s \,\mathrm{d}\sigma_s - \frac{1}{2}\int_0^T \left(\frac{\partial}{\partial s}V_s\right)^2 \mathrm{d}s} \,\mathrm{d}\gamma(\sigma)$$

$$= \int_{C_0^{T_0}\times C_0^{T-T_0}} \tilde{F}\circ \tilde{T}_V e^{-\int_0^{T_0} \dot{V}_s \,\mathrm{d}\sigma_s - \frac{1}{2}\int_0^{T_0}\left(\dot{V}_s(\sigma)\right)^2 \mathrm{d}s} e^{-\int_0^{T-T_0} \dot{\alpha}(\sigma)_s \,\mathrm{d}\rho_s - \frac{1}{2}\int_0^{T-T_0}(\dot{\alpha}_k(\sigma)_s)^2\,\mathrm{d}s} \,\mathrm{d}\gamma^{T-T_0}(\rho)\mathrm{d}\gamma^{T_0}(\sigma)$$

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

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

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

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

$$= \int_{C_0} F(\sigma)\,\mathrm{d}\gamma(\sigma)$$

as required. ◻

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

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

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 $\tau = 0$. ◻

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

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

$$\langle \nabla_H F(x), h\rangle_H = \mathrm{D}_H F(x)(h) = \lim_{t\to 0} \frac{F(x + ti(h)) - F(x)}{t}.$$

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

Remark 9.4. Lemma 9.2 can be written

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

where $\mathrm{div}\,V : C_0 \to \mathbb{R}$ is $-\int_0^T \dot{V}_s\,\mathrm{d}B_s$.

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:

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

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

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 $\overline{\operatorname{span} S}^\perp = 0 \iff \overline{\operatorname{span} S} = H$. □

Proposition 9.7. *In an* AWS *$i : H \to E$, $\left\{ e^{\langle h,-\rangle^{\sim}_H - \frac{1}{2}\|h\|_H^2} \middle| h \in H \right\}$ 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 $f : E \to \mathbb{R}$ is in L^2 with $\int_E f(x) e^{\langle h,-\rangle^{\sim}_H(x)} \,\mathrm{d}\gamma(x) = 0$ for all $h \in j(E^*)$. Taking $h = j(\ell)$ this gives that for all $\ell \in E^*$,

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

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

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

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

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

The result then follows from Lemma 9.8. □

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

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

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

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

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

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 $E \to \mathbb{R}$ are dense in $L^2(E, \mu; \mathbb{R})$.

Proposition 9.10. $\left\{ 1, \sigma \mapsto \int_0^T \alpha_s \,\mathrm{d}\sigma(s) \middle| \alpha \in \mathscr{E} \right\}$ *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

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

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

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

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

where $\mathscr{F}_t$ is the natural filtration of canonical BM, $\sigma\{\mathrm{ev}_s \,|0 \le s \le t\}$.

Proof. Set $G(\sigma) := \int_0^T \dot{V}_s(\sigma) \,\mathrm{d}\sigma(s)$; $\dot{V}$ 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 is L^2. Set $\bar{F} := F - \int_{C_0} F \,\mathrm{d}\gamma$, so $\int_{C_0} \bar{F} \,\mathrm{d}\gamma = 0$, so $\bar{F} \perp$ all constants in L^2. Then

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

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:

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

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

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

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

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

for all G. Since

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

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

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

as claimed. □

Remark 9.12. This says that for $F \in BC^1$, $F = \int F + \operatorname{div} U$ for some U, since if F is BC^1, $F = \operatorname{div} U$ 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 $V : C_0 \to L_0^{2,1}$ be such that $\dot{V} : [0,T] \times C_0 \to \mathbb{R}$ is in $L^2(B)$, and so is adapted. Let $F : C_0 \to \mathbb{R}$ be BC^1. Then*

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

i.e.

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

where $\operatorname{div} V : C_0 \to \mathbb{R}$ *is* $\operatorname{div} V(\sigma) := -\int_0^T \dot{V}_s(\sigma)\,\mathrm{d}\sigma(s)$.

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

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

Thus

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

□

The above result shows that Itō 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* $\alpha^F : [0,T] \times C_0 \to \mathbb{R}$ *in* $L^2(B)$ *such that*

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

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

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

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

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

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

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

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 $F \mapsto F - \int_{C_0} F \,\mathrm{d}\gamma$ maps BC^1 to $\hat{L}^2 \cap BC^1$. Therefore, U has a unique continuous linear extension $\bar{U} : \hat{L}^2 \to L^2(B)$ that is norm-preserving. Since $\bar{\mathscr{I}} \circ U = \mathrm{id}$ on the dense subset $\hat{L}^2 \cap BC^1$, $\bar{\mathscr{I}} \circ \bar{U} = \mathrm{id}$ on $\hat{L}^2$. So for $f \in \hat{L}^2$,

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

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

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

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 $\int_0^T \beta(\sigma)_s \,\mathrm{d}\sigma(s) = 0$. Therefore,

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

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

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

$$f = \mathbb{E}f + \int_0^T a_s^f \, \mathrm{d}B_s \text{ almost surely.}$$

Proof. Consider the map $\Omega \to C_0 : \omega \mapsto B.(\omega)$, so $\mathscr{F}_t = \sigma(B.)$. 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

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

Therefore,

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

But the composition

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

is

$$\omega \mapsto \left(\int_0^T \alpha^F(B.(-))_s \, \mathrm{d}B_s \right)(\omega)$$

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

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

$$M_t = M_0 + \int_0^t \alpha_s \, \mathrm{d}B_s \text{ almost surely}$$

for $0 \le t \le T$.

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

$$M_T = \mathbb{E}M_T + \int_0^T a_s^{M_T} \, \mathrm{d}B_s.$$

Therefore, by the martingale property of Itō integrals,

$$M_t = \mathbb{E}\{M_T|\mathscr{F}_t^B\} = \mathbb{E}M_T + \int_0^t a_s^{M_t} \, \mathrm{d}B_s.$$

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

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

Uniqueness holds just for M_T by Theorem 9.15. ☐

Remark 9.17. We claimed (but did not prove) that we can choose $\left(\int_0^t \alpha_s \, \mathrm{d}B_s\right)(\omega)$ 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 $\mathscr{F}_T$);

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

Example 9.18. Consider $\Omega = [0,1]$ with Lebesgue measure λ^1 and the process $a : [0,1] \times \Omega \to \mathbb{R}$ 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 $\mathscr{F}^B_*$-martingale with M_T in L^2, there exists a process $\langle M \rangle : [0,T] \times \Omega \to \mathbb{R}_{\geq 0}$ 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 $\mathscr{F}^B_*$-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 $\alpha \in L^2(B)$,

$$M_t = M_0 + \int_0^t \alpha_s \, \mathrm{d}B_s.$$

Apply Itō's formula

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

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

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

So

$$M_t^2 = M_0^2 + 2 \int_0^t M_s \alpha_s \, \mathrm{d}B_s + \frac{1}{2} 2 \int_0^t \alpha_s^2 \, \mathrm{d}s.$$

Hence,

$$M_t^2 - \int_0^t \alpha_s^2 \, \mathrm{d}s = M_0^2 + 2 \int_0^t M_s \alpha_s \, \mathrm{d}B_s,$$

and since $M.$ is bounded, $M.\alpha. \in L^2(B)$, and so the RHS is an $\mathscr{F}^B_*$-martingale. (M_0 is constant since it is $\mathscr{F}^B_0$-measurable, and $\mathscr{F}^B_0 = \{\emptyset, \Omega\}$.) Now set $\langle M \rangle_t := \int_0^t \alpha_s^2 \, \mathrm{d}s$. □

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

9.2 Chaos Expansions

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

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

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

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

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

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

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

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

where $\bar{\alpha}^{(k-1)} \in L^2(\{0 \le t_1 \le \dots \le t_{k-1} \le T\} \subseteq [0,T]^{k-1};\mathbb{R})$ and $\alpha^{(k)}_{t_1,\dots,t_k}$ is $\mathscr{F}_{t_1}$-measurable in $L^2(\{0 \le t_1 \le \dots \le t_k \le T\} \times \Omega; \mathbb{R})$. 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. $\mathbb{E} \int_0^T \int_0^t a_{s,t}\,\mathrm{d}\sigma(s)\mathrm{d}\sigma(t) \int_0^T b_s\,\mathrm{d}\sigma(s) = 0$. By the isometry property,

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

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