

SOBRE CONVERGENCIA DE DENSIDADES DE PROMEDIOS ESPACIALES PARA LA ECUACIÓN ESTOCÁSTICA DEL CALOR

T E S I S

Que para obtener el grado de

Maestría en Ciencias con Orientación en Probabilidad y Estadística

Presenta

Ivan Irving Rosas Domínguez

Director de Tesis:

Dr. Arturo Jaramillo Gil

Autorización de la versión final

Índice general

1. Elementos de ecuaciones diferenciales parciales estocásticas	1
1.1. Introducción a SPDEs	1
1.1.1. Heurística de la formulación Mild	4
1.2. Integral de Itô-Walsh	5
1.2.1. Variables Gaussianas y Procesos Gaussianos	5
1.2.2. Medidas martingala	12
1.2.3. Definición de la Integral	16
1.3. Formulación Mild de una SPDE. Existencia y unicidad de las soluciones	20
2. Elementos de Cálculo de Malliavin	23
2.1. Cálculo de Malliavin en el caso unidimensional	23
2.1.1. La derivada de Malliavin.	23
2.1.2. El operador de divergencia	29
2.1.3. El semigrupo de Ornstein-Uhlenbeck	31
2.1.4. El generador infinitesimal	36
2.2. Cálculo de Malliavin en el caso general.	40
3. Aplicaciones a Teoremas Límite y estudio de densidades	43
3.1. Método de Stein	43
3.2. Existencia de densidades y estimación de la distancia uniforme	50
4. Propiedades espaciales de la Ecuación estocástica del calor	55
4.1. Formulación Mild en el caso sin deriva	60
4.2. Densidad explícita de los promedios espaciales	60
4.3. Aproximación uniforme de las densidades	60

Capítulo 1

Elementos de ecuaciones diferenciales parciales estocásticas

En este capítulo hablaremos acerca del concepto de ecuación diferencial parcial estocástica. Dichos objetos comenzaron a ser estudiados a mediados del siglo pasado con la llegada de la integral de Itô. El trabajo de Itô y demás colegas para tratar ecuaciones diferenciales estocásticas inmediatamente sugirió el estudio de la posible generalización a ecuaciones diferenciales en varias variables. Durante las décadas de los 50, 60 y 70, varios artículos hacían alusión a estos objetos, aunque no fue sino hasta finales de la década de 1970 cuando finalmente estas ecuaciones comenzaron a ser estudiadas como objetos matemáticos en sí. Finalmente, en la década de 1980, varias monografías surgieron con el propósito de construir una teoría que cohesionara las ideas existentes en el área. Destaca en particular la monografía escrita por Walsh [8] y publicada en 1986, en donde el autor aborda el estudio de éstas ecuaciones, dando un significado preciso a lo que significa plantear y resolver una ecuación diferencial que involucre procesos estocásticos cuyas trayectorias generalmente no son diferenciables, y que tengan variables tanto espaciales como una temporal.

Desde esta década múltiples trabajos en esta área han sido publicados. Esto, junto con el uso de herramientas como cálculo de Malliavin en el estudio de las mismas, han hecho de las Ecuaciones diferenciales parciales estocásticas una área bastante activa en la probabilidad moderna.

1.1. Introducción a SPDEs

Comenzamos en este capítulo con la pregunta clave: ¿qué es una ecuación diferencial parcial estocástica? Para responder a esta pregunta, podemos repasar cada uno de los conceptos que nos llevan a esta idea. Partimos del mundo clásico en \mathbb{R}^N : sean $x \in \mathbb{R}^N$, $t \in [0, \infty)$ y $u : \mathbb{R}^N \times [0, \infty) \rightarrow \mathbb{R}$ una función en las variables x y t . Una ecuación diferencial parcial es una

ecuación del tipo

$$F(t, x, u, Du, D^2u, \dots) = 0, \quad (1.1)$$

donde x y t se interpretan como las variables espacial y temporal respectivamente, y F es una función arbitraria que depende tanto de las variables espaciales como de la temporal, así como de la misma función u y de sus derivadas de orden $\alpha = (\alpha_1, \dots, \alpha_n)$, donde α_i representa la derivada parcial en la i -ésima coordenada.

Si existe una función u tal que $u \in C^\alpha(\mathbb{R}^N \times [0, \infty))$ y además u satisface la ecuación (1.1), donde α es el multi-índice más grande tal que el orden de todas las derivadas que aparecen en la ecuación son cubiertas por el mismo, entonces decimos que u es una solución clásica a la ecuación diferencial parcial anterior. De esta manera, y en particular para la ecuación estocástica del calor homogénea clásica, la cual está dada por

$$\partial_t u(x, t) - \frac{1}{2} \Delta_x u(x, t) = 0,$$

se entiende por solución clásica a una función $u : \mathbb{R}^N \times [0, \infty) \rightarrow \mathbb{R}$ tal que $u \in C^{2,1}(\mathbb{R}^N \times [0, \infty))$, esto es, al menos todas las segundas derivadas en las variables espaciales existen y son continuas, y al menos la derivada en la variable temporal existe y es continua.

Cabe destacar que la noción de solución de una ecuación diferencial parcial en el sentido clásico coincide con lo esperable: una función u será solución de una ecuación diferencial si al menos tiene tantas derivadas como aquellas involucradas en la igualdad (1.1).

No obstante, en algunas ocasiones los problemas planteados requieren soluciones que no necesariamente cumplan tales características de regularidad, pero que intuitivamente deberían tener sentido. Un ejemplo muy conocido es aquél de la ecuación de onda con una condición inicial que no necesariamente es regular. Por ejemplo, si hablamos del problema homogéneo de la ecuación de onda unidimensional dado por:

$$\begin{cases} \partial_{tt} u(x, t) - \kappa^2 \partial_{xx} u(x, t) = 0 & (x, t) \in [-\pi, \pi] \times [0, \infty), \\ u(-\pi, t) = u(\pi, t) = 0, & t \in [0, \infty), \\ u(x, 0) = u_0(x) = \pi - |x| & x \in [-\pi, \pi], \\ u_t(x, 0) = 0 & (x, t) \in [-\pi, \pi] \times [0, \infty), \end{cases} \quad (1.2)$$

es claro que tenemos una inconsistencia al momento de colocar la condición inicial dada por el valor absoluto $\pi - |x|$: dicha función no es derivable en $x = 0$. No obstante, desde un punto de vista intuitivo, claramente el problema tiene sentido: estamos modelando el comportamiento de una cuerda sujeta en sus extremos a los puntos $-\pi$ y π , la cual al soltarla tiene una posición inicial dada por $\pi - |x|$ y cuya velocidad inicial es constante 0 en cualquier punto de la cuerda.

La idea para hallar una solución al problema anterior a grandes rasgos consiste en reformular el problema de forma que se conserve la esencia del problema, y posteriormente obtener una solución a este problema reformulado. Una manera estándar es utilizar la teoría de distribuciones. En el caso particular de la ecuación de onda anterior, se procede formalmente de la siguiente manera:

Supongamos que u es una solución al problema (1.2). En particular se debería cumplir que

$$\partial_{tt}u(x, t) = \kappa^2 \partial_{xx}u(x, t),$$

por lo que para cualquier función $\phi \in C_c^\infty([-\pi, \pi] \times [0, \infty))$ tal que se anule en la frontera,, se tiene que al multiplicar e integrar, la ecuación anterior se convierte en:

$$\int_{[-\pi, \pi] \times [0, \infty)} \partial_{tt}u(x, t)\phi(x, t)dxdt = \kappa^2 \int_{[-\pi, \pi] \times [0, \infty)} \partial_{xx}u(x, t)\phi(x, t)dxdt,$$

y suponiendo que es válido utilizar integración por partes, la igualdad anterior equivale a

$$\int_{[-\pi, \pi] \times [0, \infty)} u(x, t)\partial_{tt}\phi(x, t)dxdt = \kappa^2 \int_{[-\pi, \pi] \times [0, \infty)} u(x, t)\partial_{xx}\phi(x, t)dxdt,$$

en donde los términos evaluados en la frontera se anulan gracias a que las funciones ϕ (a menudo denominadas funciones de prueba) se anulan en la misma. Reordenando los términos anteriores, se obtiene una reformulación de la ecuación diferencial del problema (1.2) la cual no necesita que la función u sea derivable. A saber, decimos que u es una solución débil de la ecuación diferencial asociada al problema (1.2) si se cumple que

$$\int_{[-\pi, \pi] \times [0, \infty)} u(x, t)(\partial_{tt}\phi(x, t) - \kappa^2 \partial_{xx}\phi(x, t))dxdt = 0$$

para cualquier función $\phi \in C_c^\infty([-\pi, \pi] \times [0, \infty))$ que se anule en la frontera.

Claramente toda solución clásica (o fuerte) es a su vez una solución débil, ya que precisamente la regularidad de dichas soluciones permiten hacer de manera rigurosa los cálculos formales anteriores. No obstante, la existencia de una solución débil en el sentido anterior no garantiza que dicha solución sea regular, por lo que en principio es más sencillo demostrar la existencia o unicidad de las soluciones débiles, a pesar de que estas gocen de menos propiedades que las soluciones fuertes o clásicas, las cuales no siempre existirán, como en el caso del problema (1.2).

Llegados a este punto, reflexionamos ahora en dirección de la probabilidad. Supongamos nuevamente que tenemos la ecuación de onda unidimensional, pero considerando el caso no homogéneo. Tenemos entonces la siguiente ecuación

$$\partial_{tt}u(x, t) = \kappa^2 \partial_{xx}u(x, t) + F(x, t), \quad (x, t) \in [-\pi, \pi] \times [0, \infty) \quad (1.3)$$

donde ahora F es una función que representa la cantidad de presión por unidad de longitud que se aplica a la cuerda. Bajo condiciones de regularidad en la función F , podemos resolver

el problema de manera clásica utilizando el método de separación de variables. Sin embargo, nos preguntamos ¿qué sucede si ahora F representa una función que no sea diferenciable? Más aún, ¿qué pasa si consideramos a F una *perturbación aleatoria*?

Para dar un sentido más preciso a las ideas anteriores podemos pensar, para x y t fijos, a $F(x, t)$ como una variable aleatoria. Dicha variable depende tanto del espacio como del tiempo, y por ende podemos interpretar a F como un proceso estocástico que evoluciona de manera espacio-temporal. Tal y como se comentó al inicio, es conocido que existen procesos estocásticos cuyas trayectorias son no diferenciables, por lo que en principio pensar en resolver una ecuación diferencial de este estilo de manera clásica se vislumbra una tarea complicada.

Más aún, ¿en qué sentido se debe interpretar una ecuación como (1.3)?, o bien, considerando la siguiente ecuación del calor no homogénea

$$\partial_t u(x, t) = \frac{1}{2} \Delta_x u(x, t) + F(x, t), \quad (x, t) \in \mathbb{R}^N \times [0, \infty), \quad (1.4)$$

en donde el término F sea aleatorio, ¿bajo qué significado hablamos de una solución de dicha ecuación?

Obsérvese que, si bien la teoría de distribuciones nos permite lidiar con la parte de la no regularidad de las condiciones del problema, para replicar el mecanismo que fue utilizado para reformular (1.2) en un sentido débil, nos encontramos con una dificultad, ya que a no ser que el proceso $F(x, t)$ sea de variación finita, no será posible realizar de manera rigurosa la transición al momento de integrar por partes.

1.1.1. Heurística de la formulación Mild

Un vistazo más a detalle de lo anterior puede arrojar luz sobre cómo sortear el problema. Suponiendo que lo anterior es válido, una manera de proceder con la solución de la ecuación del calor no homogénea será la siguiente. Supongamos que tenemos la siguiente ecuación diferencial del calor unidimensional no lineal:

$$\partial_t u(x, t) = \frac{1}{2} \partial_{xx} u(x, t) + W(x, t)u(x, t), \quad (x, t) \in [0, L] \times [0, \infty)$$

donde $W(x, t)$ representa nuestra fuente de aleatoriedad (aunque bien puede pensarse como una función suficientemente suave en principio). Utilizando la técnica presentada en el ejemplo de la ecuación de onda, podríamos tomar una función $\phi \in C^\infty([0, L])$ cuya derivada se anule en la frontera, multiplicarla por la ecuación anterior e integrar con respecto al espacio y al tiempo, obteniendo que

$$\int_0^L u(x, t) \phi(x) dx - \int_0^L u(x, 0) \phi(x) dx = \int_0^L \int_0^t \partial_{xx} u(x, s) \phi(x) dx ds + \int_0^L \int_0^t \partial_{xx} f(x, s) \phi(x) W(x, s) dx ds$$

Resta entonces estudiar en qué sentido podemos integrar la ecuación (1.3) o (1.4) para sortear las dificultades que la aleatoriedad añade al problema.

1.2. Integral de Itô-Walsh

En esta sección presentamos la construcción de la integral de Itô-Walsh. Dicha integral nos permite construir procesos estocásticos que más adelante y en cierto sentido pueden verse como la solución de una ecuación diferencial parcial estocástica. Para lograr dicho objetivo, presentamos primeramente la teoría necesaria para construir la integral.

1.2.1. Variables Gaussianas y Procesos Gaussianos

Recordamos aquí las propiedades fundamentales de los procesos gaussianos. Salvo algunos casos especiales, la mayoría de las propiedades se enuncian sin demostración, por lo que nos referiremos a [6] para revisar las pruebas de las mismas. Comenzamos definiendo la distribución Gaussiana.

Definición 1. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y consideremos a los reales dotados de la σ -álgebra de Borel, la cual denotamos por $\mathcal{B}(\mathbb{R})$. Decimos que una variable aleatoria $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ tiene distribución normal estándar (o bien, es una variable Gaussiana estándar), si la función de distribución de la misma está dada por:

$$F_X(t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

En otras palabras, si la variable aleatoria X tiene densidad (con respecto a la medida de Lebesgue en \mathbb{R}) dada por la función

$$f_X(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

En tal caso, denotamos dicha situación por $X \sim \text{Normal}(0,1)$.

Directamente de la definición, tenemos la caracterización de la variable Gaussiana en términos de su transformada de Laplace y de su transformada de Fourier.

Proposición 1. Sea X una variable aleatoria con distribución normal estándar. Su transformada de Laplace en todo \mathbb{C} está dada por

$$\mathcal{L}_X(\lambda) := \mathbb{E} [e^{\lambda X}] = e^{\lambda^2/2}, \quad \lambda \in \mathbb{C},$$

mientras que su transformada de Fourier (o función característica) en \mathbb{R} está dada por

$$\mathcal{F}_X(\xi) := \mathbb{E} [e^{-i\xi X}] = e^{-\xi^2/2}, \quad \xi \in \mathbb{R}.$$

Se puede ver usando alguna de las transformadas anteriores y usando un argumento inductivo que una variable normal estándar tiene momentos n -ésimos de cualquier orden, y que los mismos están dados por

- $\mathbb{E}[X^{2n}] = (n-1)!! = \frac{(2n)!}{2^n n!} = 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1), \quad n \geq 0.$
- $\mathbb{E}[X^{2n+1}] = 0, \quad n \geq 0.$

En particular, se cumple que $\mathbb{E}[X^{n+1}] = n \cdot \mathbb{E}[X^{n-1}]$, para $n \geq 1$.

Diremos también que una variable aleatoria Gaussiana tiene media $\mu \in \mathbb{R}$ y varianza $\sigma^2 > 0$ si se cumple que $Y = \sigma X + \mu$ para X una variable normal estándar. En tal caso, denotaremos la situación por $Y \sim \text{Normal}(\mu, \sigma^2)$. En vista de la proposición anterior y de la definición, se tiene la siguiente proposición.

Proposición 2. Sea Y una variable aleatoria normal con media $\mu \in \mathbb{R}$ y varianza $\sigma^2 > 0$. Entonces las siguientes tres proposiciones son equivalentes a este hecho.

- $\mathcal{L}_Y(\lambda) = e^{\mu\lambda + \sigma^2\lambda^2/2}, \quad \lambda \in \mathbb{C},$
- $\mathcal{F}_X(\xi) = e^{i\mu\xi - \sigma^2\xi^2/2}, \quad \xi \in \mathbb{R},$
- La distribución de Y tiene densidad con respecto a la medida de Lebesgue en \mathbb{R} dada por

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)/2\sigma^2}, \quad t \in \mathbb{R}.$$

Por convención, diremos que una variable Y tiene distribución normal con media $\mu \in \mathbb{R}$ y varianza $\sigma^2 = 0$ (denotado por $Y \sim \text{Normal}(\mu, 0)$) si se cumple que $Y = \mu$, λ -casi seguramente, donde λ denota la medida de Lebesgue en \mathbb{R} .

Pasamos ahora a definir los vectores Gaussianos o vectores normales conjuntos.

Definición 2. Sea $(\mathbb{R}^d, \langle \cdot, \cdot \rangle)$ el espacio \mathbb{R}^d dotado de su producto interno usual. Decimos que una variable aleatoria $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ es un vector gaussiano, o un vector normal conjunto, si para cualquier $u \in \mathbb{R}^d$, se tiene que $\langle u, X \rangle$ es una variable aleatoria Gaussiana (en \mathbb{R}).

Presentamos un primer resultado sobre vectores aleatorios Gaussianos.

Proposición 3. Sea X un vector aleatorio Gaussiano en \mathbb{R}^d . Entonces existe una función $\mu_X \in \mathbb{R}^d$ y una forma cuadrática no negativa $q_X : \mathbb{R}^d \rightarrow \mathbb{R}$ tales que para cualquier $u \in \mathbb{R}^d$,

- $\mathbb{E}[\langle u, X \rangle] = \langle u, \mu_X \rangle$
- $\text{Var}(\langle u, X \rangle) = q_X(u).$

De hecho, si tenemos una base ortonormal $(e_j)_{j \geq 1}$ de \mathbb{R}^d , entonces podemos escribir a X de la siguiente manera:

$$X = \sum_{j=1}^d \langle X, e_j \rangle e_j,$$

donde $X_j := \langle X, e_j \rangle$, por definición, es una variable aleatoria Gaussiana, para cualquier $j \geq 1$. De lo anterior, se sigue que, escribiendo $u \in \mathbb{R}^d$ como $u = \sum_{j=1}^d u_j e_j$, se tiene que

- $\mu_X = \mathbb{E}[X_j] e_j.$
- $q_X(u) = \sum_{j=1}^d \sum_{i=1}^d u_i u_j \text{Cov}(X_i, X_j).$

Se sigue de la proposición anterior la fórmula de la función característica de un vector gaussiano.

Proposición 4. Sea X un vector aleatorio gaussiano y $(e_j)_{j \geq 1}$ una base ortonormal de \mathbb{R}^d . Entonces, al ser $\langle u, X \rangle \sim \text{Normal}(\langle u, \mu_X \rangle, q_X(u))$, se tiene que

$$\mathbb{E}[e^{i\langle u, X \rangle}] = e^{i\langle u, \mu_X \rangle - \frac{1}{2}q_X(u)}, \quad u \in \mathbb{R}^d.$$

Y en vista de las dos proposiciones anteriores, está la siguiente caracterización de variables aleatorias Gaussianas independientes.

Proposición 5. Sea X un vector aleatorio gaussiano en \mathbb{R}^d , y sea $(e_j)_{j \geq 1}$ una base ortonormal de \mathbb{R}^d . Entonces X_1, \dots, X_d son variables aleatorias independientes si y solo si la matriz de covarianzas $\Sigma = (\text{Cov}(X_i, X_j))_{1 \leq i, j \leq d}$ es diagonal. Equivalentemente, X_1, \dots, X_d son independientes si y solo si la forma cuadrática q_X está en forma diagonal en la base $(e_j)_{j \geq 1}$.

Pasamos ahora a estudiar procesos estocásticos formados por variables aleatorias gaussianas.

Definición 3. Sea T un conjunto arbitrario y sea $G = (G(t))_{t \in T}$ una colección de variables aleatorias indexadas por T . Decimos que G es un proceso Gaussiano, o un campo aleatorio Gaussiano si para cualesquiera $t_1, \dots, t_k \in T$, se tiene que $X = (G(t_1), \dots, G(t_k))$ es un vector gaussiano, $k \geq 1$.

Equivalentemente, un proceso estocástico $(G(t))_{t \in T}$ es Gaussiano si cualquier combinación lineal finita de variables $G(t)$, $t \in T$ es Gaussiana, lo cual de acuerdo a nuestra exposición, también equivale a decir que las distribuciones finito-dimensionales de un proceso Gaussiano $(G(t))_{t \in T}$ están dictadas por medio de variables aleatorias gaussianas.

A menudo es importante saber cuando un proceso gaussiano es independiente de otro proceso gaussiano. Para ello, recordamos la definición de σ -álgebra generada por un conjunto de variables.

Definición 4. Sea H una familia de variables aleatorias definidas en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. La σ -álgebra generada por H es la σ -álgebra más pequeña en Ω tal que todas las variables $X \in H$ son medibles con respecto a dicha σ -álgebra. A tal estructura la denotamos como $\sigma(H)$.

Enunciamos ahora un importante resultado con respecto a la independencia de dos conjuntos de variables aleatorias gaussianas. Este resultado nos dice que, en cierto sentido, dos conjuntos de variables aleatorias Gaussianas son ortogonales si y solo si son independientes entre sí.

Proposición 6. Sean G_1, G_2 dos familias de variables aleatorias Gaussianas centradas (i.e. con media cero). Denotemos por H_1 y H_2 al subespacio lineal cerrado de $L^2(\mathbb{P})$ generado por

dichas variables. Entonces dichos espacios están formados por variables aleatorias Gaussianas centradas.

Más aún, son equivalentes

- Los subespacios H_1 y H_2 son ortogonales entre sí en $L^2(\mathbb{P})$.
- Las σ -álgebras $\sigma(H_1)$ y $\sigma(H_2)$ son independientes.

Es importante recordar que esta propiedad es muy particular del contexto Gaussiano.

Ahora bien, dado un proceso Gaussiano G , tenemos dos funciones asociadas al mismo.

Definición 5. Sea $G = (G(t))_{t \in T}$ un proceso gaussiano. Definimos las funciones de media y covarianza del proceso G respectivamente como sigue.

1. $\mu(t) := \mathbb{E}[G(t)]$, para cualquier $t \in T$,
2. $\Gamma(s, t) := \text{Cov}(G(s), G(t))$, para cualesquiera $s, t \in T$.

Dichas funciones con valores en \mathbb{R} determina la colección de distribuciones finito-dimensionales del proceso. De hecho, gracias a las proposiciones anteriores es claro que para cualesquiera subíndices $(t_1, \dots, t_k) \subseteq T$, la ley del vector Gaussiano $X = (G(t_1), \dots, G(t_k))$ está determinada de manera única.

Lo anterior pues $\mu_X := (\mathbb{E}[G(t_1)], \dots, \mathbb{E}[G(t_k)]) = (\mu(t_1), \dots, \mu(t_k))$ y q_X la forma cuadrática asociada que está determinada por la matriz de covarianzas de X , está en términos de la función Γ definida antes, a saber, $\Sigma_X = (\text{Cov}(G(t_i), G(t_j)))_{1 \leq i, j \leq k} = (\Gamma(t_i, t_j))_{1 \leq i, j \leq k}$.

Es consecuencia de que la forma cuadrática q_X para cualquier vector X que componga una distribución finito-dimensional de G , sea no negativa definida y simétrica, el hecho de que la función de covarianzas Γ sea no negativa definida. A saber, si $c : T \rightarrow \mathbb{R}$ es una función con soporte finito, entonces

$$\sum_{T \times T} c(s)c(t)\Gamma(s, t) \geq 0$$

Lo valioso de conocer las funciones anteriores es que, al caracterizar completamente las distribuciones finito-dimensionales del proceso Gaussiano G , todo el proceso está caracterizado por medio de dichas funciones, por lo que basta conocer dichas funciones para construir un proceso Gaussiano. Este es el contenido del siguiente teorema.

Proposición 7. Sean $\Gamma : T \times T \rightarrow \mathbb{R}$ y $\mu : T \rightarrow \mathbb{R}$, tales que Γ es simétrica y no negativa definida. Entonces existe un proceso gaussiano $(G(t))_{t \in T}$ en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ apropiado, tal que para cualesquiera $t \in T$, $G(t) \sim \text{Normal}(\mu(t), \Gamma(t, t))$, con μ y Γ las respectivas funciones de media y covarianza del proceso G .

La demostración de lo anterior es una consecuencia del Teorema de Consistencia (o extensión) de Kolmogorov (ver Teorema 6.3 en [6]). Para terminar esta sección, a continuación presentamos un par de ejemplos de procesos Gaussianos contruidos de esta forma.

Ejemplo 1 (El movimiento Browniano). Consideremos el caso en el que $T = [0, \infty)$, $\mu(t) = 0$ para cualquier $t \geq 0$ y $\Gamma(s, t) = s \wedge t = \min\{s, t\}$, para $s, t \geq 0$. Claramente Γ es simétrica, y además nótese que para cualesquiera $c : T \rightarrow \mathbb{R}$ función de soporte finito, se tiene que

$$\begin{aligned}
\sum_{T \times T} c(s)c(t)\Gamma(s, t) &= \sum_{s \in T} \sum_{t \in T} c(s)c(t)(s \wedge t) \\
&= \sum_{s \in T} \sum_{t \in T} c(s)c(t) \int_0^\infty \mathbf{1}_{[0, s]}(x) \mathbf{1}_{[0, t]}(x) dx \\
&= \int_0^\infty \sum_{s \in T} \sum_{t \in T} \mathbf{1}_{[0, s]}(x) c(s) \mathbf{1}_{[0, t]}(x) c(t) dx \\
&= \int_0^\infty \left| \sum_{t \in T} \mathbf{1}_{[0, t]}(x) c(t) \right|^2 dx \\
&\geq 0,
\end{aligned}$$

por lo que por la proposición anterior, existe en un espacio de probabilidad adecuado un proceso gaussiano que denotaremos por $(B(t))_{t \geq 0}$, con función de medias 0 y función de covarianza $\Gamma(s, t) = s \wedge t$. Dicho proceso es el movimiento Browniano estándar.

El siguiente ejemplo es de suma importancia y es parte esencial en el resto del texto.

Ejemplo 2 (El ruido blanco en \mathbb{R}^d). Sea $T = \mathcal{B}(\mathbb{R}^d)$ el conjunto de los borelianos en \mathbb{R}^d . Dado que estamos hablando de subconjuntos de \mathbb{R}^d , cambiaremos la notación al indicar con letras mayúsculas a los elementos de T . Consideremos nuevamente la función de medias $\mu(A) := 0$, para cualquier $A \in \mathcal{B}(\mathbb{R}^d)$ y ahora consideremos la función $\Gamma(A, B) := \lambda^d(A \cap B)$, donde $\lambda^d : \mathcal{B}(\mathbb{R}^d) \rightarrow [0, 1]$ es la medida de Lebesgue en \mathbb{R}^d .

Claramente Γ es una función simétrica, y además, para $c : T \rightarrow \mathbb{R}$ función de soporte finito, se tiene que

$$\begin{aligned}
\sum_{T \times T} c(A)c(B)\Gamma(A, B) &= \sum_{A \in T} \sum_{B \in T} c(A)c(B)\lambda^d(A \cap B) \\
&= \sum_{A \in T} \sum_{B \in T} c(A)c(B) \int_{\mathbb{R}^d} \mathbf{1}_A(x) \mathbf{1}_B(x) \lambda^d(dx) \\
&= \int_{\mathbb{R}^d} \sum_{A \in T} \sum_{B \in T} \mathbf{1}_A(x) c(A) \mathbf{1}_B(x) c(B) \lambda^d(dx) \\
&= \int_{\mathbb{R}^d} \left| \sum_{A \in T} \mathbf{1}_A(x) c(A) \right|^2 \lambda^d(dx) \\
&\geq 0,
\end{aligned}$$

por lo que existe un proceso Gaussiano que denotaremos por $(\dot{W}(A))_{A \in \mathcal{B}(\mathbb{R}^d)}$ en un espacio

de probabilidad adecuado, tal que su función de medias es 0 y su función de covarianzas es $\Gamma(A, B) = \lambda^d(A \cap B)$. Dicho proceso estocástico es conocido como *ruido blanco* en \mathbb{R}^d .

Observemos que, si $A \cap B = \emptyset$, entonces $\Gamma(A, B) = \text{Cov}(\dot{W}(A), \dot{W}(B)) = 0$, por lo que al ser variables Gaussianas, estas son independientes según lo visto antes.

Se sigue que si $A, B \in \mathcal{B}(\mathbb{R}^d)$, entonces aprovechando que el proceso es Gaussiano centrado, tenemos que

$$\begin{aligned}
 & \text{Cov} \left(\dot{W}(A \cup B) - \dot{W}(A) - \dot{W}(B) + \dot{W}(A \cap B), \dot{W}(A \cup B) - \dot{W}(A) - \dot{W}(B) + \dot{W}(A \cap B) \right) \\
 &= \text{Cov} \left(\dot{W}(A \cup B), \dot{W}(A \cup B) \right) + \text{Cov} \left(\dot{W}(A), \dot{W}(A) \right) + \text{Cov} \left(\dot{W}(B), \dot{W}(B) \right) \\
 & \quad + \text{Cov} \left(\dot{W}(A \cap B), \dot{W}(A \cap B) \right) - 2\text{Cov} \left(\dot{W}(A \cup B), \dot{W}(A) \right) - 2\text{Cov} \left(\dot{W}(A \cup B), \dot{W}(B) \right) \\
 & \quad + 2\text{Cov} \left(\dot{W}(A \cup B), \dot{W}(A \cap B) \right) + 2\text{Cov} \left(\dot{W}(A), \dot{W}(B) \right) - 2\text{Cov} \left(\dot{W}(A), \dot{W}(A \cap B) \right) \\
 & \quad - 2\text{Cov} \left(\dot{W}(B), \dot{W}(A \cap B) \right) \\
 &= \lambda^d(A \cup B) + \lambda^d(A) + \lambda^d(B) + \lambda^d(A \cap B) - 2\lambda^d((A \cup B) \cap A) - 2\lambda^d((A \cup B) \cap B) \\
 & \quad + 2\lambda^d((A \cup B) \cap (A \cap B)) + 2\lambda^d(A \cap B) - 2\lambda^d(A \cap (A \cap B)) - 2\lambda^d(B \cap (A \cap B)) \\
 &= \lambda^d(A \cup B) + \lambda^d(A) + \lambda^d(B) + \lambda^d(A \cap B) - 2\lambda^d(A) - 2\lambda^d(B) + 2\lambda^d(A \cap B) \\
 & \quad + 2\lambda^d(A \cap B) - 2\lambda^d(A \cap B) - 2\lambda^d(A \cap B) \\
 &= \lambda^d(A \cup B) - \lambda^d(A) - \lambda^d(B) + \lambda^d(A \cap B) \\
 &= \lambda^d(A \cup B) - \lambda^d(A \cup B) \\
 &= 0.
 \end{aligned}$$

por lo que la variable $\dot{W}(A \cup B) - (\dot{W}(A) + \dot{W}(B) - \dot{W}(A \cap B))$ tiene varianza cero. Luego,

$$\dot{W}(A \cup B) = \dot{W}(A) + \dot{W}(B) - \dot{W}(A \cap B) \quad \mathbb{P} - \text{casi seguramente.} \quad (1.5)$$

Lo anterior nos podría hacer pensar que el ruido blanco es una medida signada. Sin embargo, esto es falso (ver ejemplo 3.16 de [4]). No obstante, podemos utilizar dicho proceso como base para construir una integral.

Ejemplo 3 (El proceso isonormal). Dado un ruido blanco $(\dot{W}(A))_{A \in \mathcal{B}(\mathbb{R}^d)}$, nos gustaría definir $\dot{W}(h)$ para alguna función h adecuada. Es bien conocido en teoría de la medida un mecanismo estándar para definir nuevos objetos a partir de objetos más sencillos, y luego extender por aproximación. Seguiremos esta maquinaria estándar a continuación.

Sea $A \in \mathcal{B}(\mathbb{R}^d)$. Definimos $\dot{W}(\mathbb{1}_A) := \dot{W}(A)$, y para cualesquiera $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R}^d)$ disjuntos, y constantes $c_1, \dots, c_n \in \mathbb{R}$, definimos

$$\dot{W} \left(\sum_{k=1}^n c_k \mathbb{1}_{A_k} \right) := \sum_{k=1}^n c_k \dot{W}(A_k)$$

Nótese que si existen dos representaciones distintas de una función simple, de acuerdo a (1.5), la variable de la definición anterior es consistente ya que habrá una igualdad \mathbb{P} - casi seguramente. Por otro lado, dado que los conjuntos A_1, \dots, A_n son disjuntos, las variables $\dot{W}(A_1), \dots, \dot{W}(A_n)$ son independientes entre sí. Luego, observamos que al ser centradas,

$$\begin{aligned}
\left\| \dot{W} \left(\sum_{k=1}^n c_k \mathbb{1}_{A_k} \right) \right\|_{L^2(\mathbb{P})}^2 &= \text{Var} \left(\sum_{k=1}^n c_k \dot{W}(A_k) \right) \\
&= \sum_{k=1}^n c_k^2 \mathbb{E} [\dot{W}^2(A_k)] \\
&= \sum_{k=1}^n c_k^2 \Gamma(A_k, A_k) \\
&= \sum_{k=1}^n c_k^2 \lambda^d(A_k) \\
&= \int_{\mathbb{R}^d} \sum_{k=1}^n c_k^2 \mathbb{1}_{A_k} \lambda^d(dx) \\
&= \left\| \sum_{k=1}^n c_k \mathbb{1}_{A_k} \right\|_{L^2(\mathbb{R}^d)}^2,
\end{aligned}$$

donde en la última igualdad utilizamos que, al ser $A_i \cap A_j = \emptyset$, se tiene que $\mathbb{1}_{A_i} \mathbb{1}_{A_j} = 0$ para $i \neq j$. De lo anterior deducimos que la aplicación $f \mapsto \dot{W}(f)$, que envía funciones simples en variables Gaussianas, preserva la norma de los respectivos espacios.

Es un resultado conocido en teoría de la medida que, si tenemos ahora una función $f \in L^2(\mathbb{R}^d)$, existe una sucesión de variables aleatorias simples en $L^2(\mathbb{R}^d)$ tales que $f_n \rightarrow f$ en la norma de dicho espacio. Luego, la sucesión $(f_n)_{n \geq 1}$ forma una sucesión de Cauchy de funciones simples en $L^2(\mathbb{R}^d)$. Dado que la norma se preserva para estas funciones entre los espacios L^2 según lo visto antes, la sucesión de variables $(\dot{W}(f_n))_{n \geq 1}$ es de Cauchy en $L^2(\mathbb{P})$, por lo que al ser un espacio completo, existe una única variable en $L^2(\mathbb{P})$, que denotaremos por $\dot{W}(f)$, tal que $\dot{W}(f_n) \rightarrow \dot{W}(f)$ en norma $L^2(\mathbb{P})$.

Consideramos así al espacio $L^2(\mathbb{R}^d)$, y a la colección de variables aleatorias $(\dot{W}(f))_{f \in L^2(\mathbb{R}^d)}$. A dicho proceso estocástico se le conoce como el *proceso isonormal*. Una característica fundamental de dicho proceso es que la aplicación $A \mapsto \dot{W}(A)$ es una isometría, conocida como *isometría de Wiener*. Al elemento $\dot{W}(f)$ para $f \in L^2(\mathbb{R}^d)$ se le conoce como *integral de Wiener de f* , y a menudo se denota como

$$\dot{W}(f) = \int f W(dx).$$

El proceso isonormal se puede también obtener vía su función de medias y su función de covarianzas. En efecto, tomando $T = L^2(\mathbb{R}^d)$, haciendo nuevamente $\mu(f) = 0$ para cualquier

$f \in L^2(\mathbb{R})$, y definiendo la función de covarianzas Γ como

$$\Gamma(f, g) = \int_{\mathbb{R}^d} fg \lambda^d(dx) = \langle f, g \rangle_{L^2(\mathbb{R}^d)}, \quad f, g \in L^2(\mathbb{R}^d),$$

entonces el proceso Gaussiano resultante, que denotamos por $(\dot{W}(f))_{f \in L^2(\mathbb{R}^d)}$ es el proceso isonormal que construimos antes vía un ruido blanco.

1.2.2. Medidas martingala

En esta parte del texto, estudiamos el concepto de medidas martingala. Dichos procesos estocásticos son una pérdida angular en la construcción de la integral de Itô-Walsh, como veremos a continuación.

Recordemos que si tenemos un ruido blanco $(W(A))_{A \in \mathcal{B}(\mathbb{R}^d)}$, entonces para cualesquiera $A, B \in \mathcal{B}(\mathbb{R}^d)$, se tiene la siguiente igualdad

$$\dot{W}(A \cup B) = \dot{W}(A) + \dot{W}(B) - \dot{W}(A \cap B), \quad \mathbb{P} - \text{casi seguramente.}$$

Sin embargo, como también se mencionó en la subsección pasada, a pesar de que esto nos puede hacer pensar en que el ruido blanco constituye una medida aleatoria signada, lo anterior es falso. No obstante, se tiene el siguiente resultado con respecto al mismo.

Proposición 8. El proceso $(\dot{W}(A))_{A \in \mathbb{R}^d}$ cumple las siguientes propiedades

- $\dot{W}(\emptyset) = 0$, $\mathbb{P} - \text{casi seguramente,}$
- Para cualquier sucesión de conjuntos ajenos $(A_n)_{n \geq 1} \subseteq \mathcal{B}(\mathbb{R}^d)$, se cumple que

$$\dot{W}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \dot{W}(A_k), \quad \mathbb{P} - \text{casi seguramente,}$$

en donde la suma infinita anterior converge en $L^2(\mathbb{P})$. En otras palabras, el ruido blanco es una medida signada, σ -finita, $L^2(\mathbb{P})$ -valuada.

Demostración. Es un resultado conocido de teoría de la medida que, si queremos probar que la aplicación anterior es una medida signada, y ya hemos probado que es finito-aditiva, entonces basta con tomarnos una sucesión decreciente de conjuntos $(A_n)_{n \geq 1} \subseteq \mathcal{B}(\mathbb{R}^d)$ tal que al intersecar todos los elementos obtengamos el conjunto vacío, y probar que $\dot{W}(A_n) \rightarrow 0$ en $L^2(\mathbb{P})$ para concluir.

En efecto, sea $(A_n)_{n \geq 1}$ una sucesión como la descrita antes. Directamente por definición del ruido blanco,

$$\left\| \dot{W}(A_n) \right\|_{L^2(\mathbb{P})}^2 = \mathbb{E} \left[\dot{W}^2(A_n) \right] = \Gamma(A_n, A_n) = \lambda^d(A_n) \xrightarrow{n \rightarrow \infty} 0,$$

ya que al ser λ^d una medida, se cumple la propiedad de continuidad con la sucesión $(A_n)_{n \geq 1}$.

Para concluir que la medida es σ -finita, dado que \mathbb{R}^d puede ser cubierto con una sucesión numerable de conjuntos compactos, basta ver que para cualquier conjunto compacto $K \subseteq \mathbb{R}^d$, se tiene que $\left\| \dot{W}(K) \right\|_{L^2(\mathbb{R}^d)}^2 < \infty$. Pero esto es claro, ya que

$$\left\| \dot{W}(K) \right\|_{L^2(\mathbb{R}^d)}^2 = \mathbb{E} \left[\dot{W}^2(K) \right] = \lambda^d(K) < \infty,$$

ya que los compactos son acotados en \mathbb{R}^d . ■

Observación 1. Es importante distinguir el que el ruido blanco forme una medida $L^2(\mathbb{P})$ -valuada a que forme una medida signada aleatoria en los borelianos de \mathbb{R}^d . Esto es, nótese que la segunda propiedad de la proposición posee el cuantificador fuera del conjunto de probabilidad 1. Más específicamente, la proposición anterior *no* está asegurando que

$$\mathbb{P} \left(\left\{ \forall (A_n)_{n \geq 1} \subseteq \mathcal{B}(\mathbb{R}^d), \quad \dot{W} \left(\bigcup_{k=1}^{\infty} A_k \right) = \sum_{k=1}^{\infty} \dot{W}(A_k) \right\} \right) = 1$$

Ahora, a menudo es posible considerar una dimensión como la dimensión temporal, de forma que se puede estudiar un ruido blanco definido en el espacio $[0, \infty) \times \mathbb{R}^d$. El hecho de pensar al intervalo $[0, \infty)$ como una dimensión temporal nos permite definir un nuevo proceso conocido como el *proceso de ruido blanco*, $(W_t)_{t \geq 0}$, definido como

$$W_t(A) := \dot{W}([0, t] \times A), \quad A \in \mathcal{B}(\mathbb{R}^d).$$

Este es un proceso estocástico que 'evoluciona en el tiempo', pero a su vez, es un ruido blanco en A . Recordemos también que la noción de filtración como sucesión de σ -álgebras ordenadas con respecto a la contención concreta la idea de proceso que evoluciona en el tiempo. Definimos entonces la filtración del proceso de ruido blanco como sigue

$$\mathcal{F}_t := \sigma \left(\dot{W}([0, s] \times A) : 0 \leq s \leq t, A \in \mathcal{B}(\mathbb{R}^d) \right), \quad t \geq 0.$$

Claramente la familia $(\mathcal{F}_t)_{t \geq 0}$ forma una filtración, y además el proceso $(W_t)_{t \geq 0}$ es adaptado a dicha filtración.

Una propiedad muy interesante que posee el proceso $(W_t)_{t \geq 0}$, visto como un proceso que evoluciona en el tiempo, es que es una medida aleatoria (en el sentido visto antes), y también una martingala.

Proposición 9. Sea $(W_t(A))_{t \geq 0, A \in \mathcal{B}(\mathbb{R}^d)}$ un ruido blanco en $[0, \infty) \times \mathbb{R}^d$. Entonces dicho proceso forma una 'medida martingala' en el siguiente sentido:

- Para cualquier $A \in \mathcal{B}(\mathbb{R}^d)$, se cumple $W_0(A) = 0$, \mathbb{P} -casi seguramente.

- Si $t > 0$, entonces W_t es una medida signada σ -finita, $L^2(\mathbb{P})$ -valuada.
- Para cualquier $A \in \mathbb{R}^d$, el proceso $(W_t(A))_{t \geq 0, A \in \mathcal{B}(\mathbb{R})}$ es una martingala de media cero.

Demostración. Nótese que por definición, $W_0(A) = \dot{W}(\{0\} \times A)$, y dicha variable es tal que

$$\mathbb{E} \left[\dot{W}^2(\{0\} \times A) \right] = \Gamma(0 \times A, 0 \times A) = \lambda_0 \cdot \lambda^d(A) = 0,$$

por lo que dicha variable aleatoria es la variable 0 con probabilidad 1.

Por otro lado, para $t > 0$, el que $W_t : \mathcal{B}(\mathbb{R}^d) \rightarrow L^2(\mathbb{P})$ sea una medida signada σ -finita se sigue de la propiedad siguiente: para cualquier $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\mathbb{E} \left[\dot{W}([0, t] \times A) \right] = \Gamma([0, t] \times A, [0, t] \times A) = \lambda([0, t]) \cdot \lambda^d(A) = t\lambda^d(A).$$

Por lo tanto, si $A = \emptyset$, claramente $\dot{W}_t(\emptyset) = 0$ con probabilidad 1, ya que $\Gamma([0, t] \times \emptyset, [0, t] \times \emptyset) = 0$. Ahora, para mostrar la σ -aditividad y la σ -finitud se procede de manera análoga a lo hecho en la proposición 8: primero, se prueba que la aplicación $W_t : \mathcal{B}(\mathbb{R}) \rightarrow L^2(\mathbb{P})$ es finito aditiva, y posteriormente se generaliza. Para la parte de la σ -finitud se procede igual.

Finalmente, para mostrar que para cualquier $A \in \mathcal{B}(\mathbb{R}^d)$ el proceso $t \rightarrow W_t(A)$ es una martingala con respecto a la filtración $(\mathcal{F}_t)_{t \geq 0}$, observamos que el proceso claramente es integrable (lo componen variables Gaussianas) y es adaptado a la filtración por definición. Resta probar que para $A \in \mathcal{B}(\mathbb{R})$ fijo,

$$\mathbb{E} [W_{s+t}(A) | \mathcal{F}_t] = W_t(A), \quad s, t \geq 0$$

Para ello haremos uso de la proposición 6. Consideramos a la familia de variables aleatorias $C_1 = \{W_u(A) : u \leq t\}$ y $C_2 = \{W_{t+s} - W_t(A) : s > 0\}$, para $t \geq 0$. Notemos que para $0 \leq u \leq t$,

$$\begin{aligned} \mathbb{E} [(W_{t+s}(A) - W_t(A))W_u(A)] &= \Gamma([0, t+s] \times A, [0, u] \times A) - \Gamma([0, t] \times A, [0, u] \times A) \\ &= \lambda([0, t+s] \cap [0, u]) \cdot \lambda^d(A) - \lambda([0, t] \cap [0, u]) \cdot \lambda^d(A) \\ &= \lambda([0, u]) \cdot \lambda^d(A) - \lambda([0, u]) \cdot \lambda^d(A) \\ &= 0. \end{aligned}$$

Por lo tanto, para cualesquiera dos variables ξ_1 y ξ_2 en C_1 y C_2 respectivamente, se tiene que $\langle \xi_1, \xi_2 \rangle_{L^2(\mathbb{P})} = \mathbb{E} [\xi_1 \xi_2] = 0$, esto es, son ortogonales. Por lo tanto, si H_1 y H_2 representan la cerradura del subespacio lineal generado por las familias C_1 y C_2 respectivamente, dichos subespacios son ortogonales.

Por lo tanto, por la proposición mencionada antes, las σ -álgebras generadas por las familias C_1 y C_2 son independientes. En particular, para $s, t \geq 0$, $\mathcal{F}_t \subseteq \sigma(C_1)$, por lo que al ser $W_{t+s} - W_t(t)$ una variable $\sigma(C_2)$ medible, se sigue que

$$\mathbb{E} [W_{t+s}(A) | \mathcal{F}_t] = \mathbb{E} [W_{t+s}(A) - W_t(A) + W_t(A) | \mathcal{F}_t] = \mathbb{E} [W_{t+s}(A) - W_t(A)] + W_t(A) = W_t(A),$$

por lo que la propiedad de martingala se satisface. ■

El proceso de ruido blanco es entonces un ejemplo de medida martingala, no obstante, este concepto se puede definir de manera mucho más general, como vemos a continuación.

Definición 6. Sea $(\mathcal{F}_t)_{t \geq 0}$ una sucesión de σ -álgebras continuas por derecha. Un proceso $(M_t(A))_{t \geq 0, A \in \mathcal{B}(\mathbb{R}^d)}$ es una medida martingala con respecto a dicha filtración si

- $M_0(A) = 0$ \mathbb{P} - casi seguramente.
- Si $t > 0$, entonces M_t es una medida signada, σ -finita, $L^2(\mathbb{P})$ -valuada.
- Para cualquier $A \in \mathcal{B}(\mathbb{R}^d)$, se tiene que $(M_t(A))_{t \geq 0}$ es una martingala con respecto a la filtración $(\mathcal{F}_t)_{t \geq 0}$ de media 0.

Por construcción, el proceso del ruido blanco es un primer ejemplo de medida martingala. La razón de introducir este concepto, es que este conjunto de procesos forman una clase adecuada de integradores. A continuación, definimos un concepto relacionado íntimamente con las medidas martingala, y que es en cierto sentido una generalización del proceso de covariación para dos martingalas en el caso unidimensional.

Definición 7. Sea M una medida martingala. El funcional de covarianza de M se define como

$$\overline{Q}_t(A, B) := \langle M_t(A), M_t(B) \rangle_t, \quad t \geq 0, A, B \in \mathcal{B}(\mathbb{R}^d).$$

Dado que para M una medida martingala, si A y $B \in \mathcal{B}(\mathbb{R}^d)$ se tiene que $(M_t(A))_{t \geq 0}$ y $(M_t(B))_{t \geq 0}$ son martingalas, entonces el proceso de covariación de ambas martingalas evaluado en $t \geq 0$ coincide con el funcional de covarianza evaluado en los borelianos A y B , y en $t \geq 0$.

Gracias a lo anterior, se deducen directamente de las propiedades de la covariación entre dos martingalas las siguientes propiedades del funcional de covarianza de una medida martingala.

Proposición 10. Sea M una medida martingala y denotemos por \overline{Q} a su funcional de covarianza. Entonces \overline{Q} cumple que

- $\overline{Q}_t(A, B) = \overline{Q}_t(B, A)$, \mathbb{P} - casi seguramente.
- Si $B \cap C = \emptyset$, entonces $\overline{Q}_t(A, B \cup C) = \overline{Q}_t(A, B) + \overline{Q}_t(A, C)$, \mathbb{P} - casi seguramente.
- $|\overline{Q}_t(A, B)|^2 \leq \overline{Q}_t(A, A)\overline{Q}_t(B, B)$ \mathbb{P} -casi seguramente, y
- $t \mapsto \overline{Q}_t(A, A)$ es \mathbb{P} - casi seguramente no decreciente.

Con el concepto de funcional de covarianza, podemos definir una función aleatoria de conjuntos en $\mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(\mathbb{R}^d) \times \mathcal{B}([0, \infty))$, que denotaremos Q , como sigue:

$$Q(A \times B \times (s, t]) := \overline{Q}_t(A, B) - \overline{Q}_s(A, B), \quad A, B \in \mathcal{B}(\mathbb{R}^d), t \geq s \geq 0$$

Y extendemos dicha función a uniones disjuntas de elementos (rectángulos) $(A_i \times B_i \times (s_i, t_i])$

para $1 \leq i \leq m$ y $A_i, B_i \in \mathcal{B}(\mathbb{R}^d)$, $(s_i, t_i] \in \mathcal{B}(\mathbb{R}^d)$ como

$$Q\left(\bigcup_{i=1}^n (A_i \times B_i \times (s_i, t_i])\right) := \sum_{i=1}^n Q(A_i \times B_i \times (s_i, t_i]).$$

Nuestro objetivo es extender la función aleatoria de conjuntos Q a conjuntos más generales. El problema es que en general no es posible hacerlo. No obstante, con estas dos herramientas teóricas, podemos definir el concepto de *worthiness* de una medida martingala, introducido por Walsh en [8]. Dicho concepto es de carácter técnico, pero será fundamental en la construcción de la integral de Itô-Walsh al permitirnos extender la definición de Q a conjuntos más generales.

Definición 8. Sea M una medida martingala. Decimos que M es *worthy* (digna) si existe una medida σ -finita $K : \mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(\mathbb{R}^d) \times \mathcal{B}([0, \infty)) \times \Omega \rightarrow [0, 1]$ tal que para cualesquiera $A, B \in \mathcal{B}(\mathbb{R}^d)$, $C \in \mathcal{B}([0, \infty))$ y $\omega \in \Omega$,

- $A \times B \mapsto K(A \times B \times C, \omega)$ es no negativa definida y simétrica,
- $\{K(A \times B \times (0, t])\}_{t \geq 0}$ es un proceso predecible para cualesquiera $A, B \in \mathcal{B}(\mathbb{R}^d)$,
- Para cualesquiera dos conjuntos compactos $A, B \in \mathcal{B}(\mathbb{R}^d)$, y $t > 0$,

$$\mathbb{E}[K(A \times B \times (0, t])] < \infty,$$

- Para cualesquiera $A, B \in \mathcal{B}(\mathbb{R}^d)$ y $t > 0$,

$$|Q(A \times B \times (0, t])| \leq K(A \times B \times (0, t]) \quad \mathbb{P} - \text{ casi seguramente.}$$

Como es usual en probabilidad, la dependencia en ω es omitida en la notación. Cuando dicha medida K existe, decimos que es una *medida dominante* para M .

1.2.3. Definición de la Integral

En este punto tenemos suficiente teoría para definir la integral de Itô-Walsh con respecto a medidas martingala. Conforme vayamos desarrollando la construcción, vamos a ir notando cierta familiaridad con la construcción de la integral de Lebesgue y la integral de Itô.

Definición 9. Una función $f : \mathbb{R}^d \times [0, \infty) \times \Omega \rightarrow \mathbb{R}$ se dice que es *elemental* si es de la forma

$$f(x, t, \omega) = X(\omega) \mathbb{1}_{(a, b]}(t) \mathbb{1}_A(x)$$

donde X es una variable aleatoria acotada y \mathcal{F}_a -medible, y además $A \in \mathcal{B}(\mathbb{R}^d)$.

Definimos la integral de Itô-Walsh para una función elemental de la forma anterior como

$$(f \cdot M)_t(B)(\omega) := X(\omega) [M_{t \wedge b}(A \cap B) - M_{t \wedge a}(A \cap B)](\omega)$$

Como una primera propiedad, tenemos la siguiente proposición

Proposición 11. Sea f una función elemental y M una medida martingala. Entonces $(f \cdot M)$ es nuevamente una medida martingala. En particular, la integral de funciones elementales con respecto a martingalas se vuelve una forma de construir nuevas medidas martingala.

Demostración. Pendiente ■

Ejemplo 4. Sea \dot{W} un ruido blanco en $\mathbb{R}^d \times [0, \infty)$ y consideremos a $(W_t(A))_{t \geq 0, A \in \mathcal{B}(\mathbb{R}^d)}$ el proceso de ruido blanco. Ya vimos que dicho proceso forma una medida martingala. Sea f una función elemental. Entonces $(f \cdot W)$ es una medida martingala.

El siguiente paso en la construcción de la integral es extender la definición a funciones simples.

Definición 10. Decimos que una función f es una función simple si existen c_1, \dots, c_n constantes reales, y f_1, \dots, f_n funciones elementales, tales que

$$f = \sum_{k=1}^n c_k f_k.$$

Denotaremos por \mathcal{S} el conjunto de las funciones simples. Definimos la integral de Itô-Walsh para funciones simples como

$$(f \cdot M)_t(B) := \sum_{j=1}^k c_j (f_j \cdot M)_t(B)$$

Proposición 12. La definición anterior es consistente, esto es, no depende de la representación de f en términos de funciones simples.

Resulta ser que una clase adecuada de funciones integradoras son las funciones f que son predecibles. Esto es, aquella σ -álgebra que es generada por todas las funciones simples. Dicha σ -álgebra la denotamos por \mathcal{P}

En este punto, deseamos extender la definición de la integral para funciones más generales, definidas sobre conjuntos no necesariamente rectangulares. Es en este punto en donde entra en juego la función Q . Dicha función es crucial, según se ve en el siguiente

Proposición 13. Sea $f \in \mathcal{S}$ y supongamos que M es una medida martingala digna. Entonces

$$\mathbb{E} \left[((f \cdot M)_t(B))^2 \right] = \mathbb{E} \left[\int_{B \times B \times (0, t]} f(x, t) f(y, t) Q(dx, dy, dt) \right].$$

Demostración. Primero debemos de hacer la demostración para funciones elementales, y posteriormente para funciones simples. ■

Nótese en la proposición anterior que estamos integrando sobre un rectángulo del estilo $B \times B \times (0, t]$, y la idea es extender la integral a regiones más generales. Para ello, necesitamos extender la definición de Q a dichas regiones, tarea que como anticipamos antes, es posible gracias al concepto de martingala digna.

De hecho, si M es una martingala digna, Q_M puede ser extendida a una medida en todo $\mathcal{B}(\mathbb{R}^d) \times \mathcal{B}^{\mathbb{R}^d} \times [0, \infty)$.

Ejemplo 5. Sea \dot{W} un ruido blanco en $\mathbb{R}^d \times [0, \infty)$ y consideremos al proceso de ruido blanco asociado al mismo. Entonces dicho proceso forma una medida martingala digna con medida dominante $K(A \times B \times C) := \lambda^d(A \cap B)\lambda(C)$.

Tenemos a continuación una proposición con respecto a las medidas martingala dignas y la integral de Itô-Walsh de funciones simples.

Proposición 14. Sea M una medida martingala digna y f una función simple. Entonces $(f \cdot M)$ es una medida martingala digna. Mas aún, si Q_N y K_N representan al funcional de covarianza y a la medida dominante de una medida martingala digna N , entonces

$$\begin{aligned} Q_{f \cdot M}(dx \, dy \, dz) &= f(x, t)f(y, t)Q_M(dx \, dy \, dt) \\ K_{f \cdot M}(dx \, dy \, dt) &= |f(x, t)f(y, t)| K_M(dx \, dy \, dt). \end{aligned}$$

Demostración. La prueba nuevamente va primero por funciones elementales y luego por funciones simples. ■

A partir de ahora, solo estaremos interesados en el caso cuando la variable temporal está en algún intervalo finito $(0, T]$. Siguiendo la maquinaria estándar para definir nuevos objetos a partir de aproximaciones, definimos ahora una norma apropiada para dicho propósito.

Definición 11. Sea M una medida martingala digna. Supongamos que K_M es su medida dominante. Sea $f \in \mathcal{P}$ una función en el conjunto de las funciones que son predecibles (es decir, que son medibles con respecto a la σ -álgebra \mathcal{P} generada por las funciones simples). Definimos la norma $\|\cdot\|_M$ como sigue

$$\|f\|_M^2 := \mathbb{E} \left[\int_{\mathbb{R}^d \times \mathbb{R}^d \times (0, T]} |f(x, t)f(y, t)| K_M(dx \, dy \, dt) \right]$$

Observamos que dicha norma está bien definida. Denotamos por \mathcal{P}_M al conjunto de todas las funciones f predecibles tales que $\mathbb{E}[\|f\|_M] < \infty$.

Proposición 15. La norma anteriormente definida en efecto es una norma, y el conjunto \mathcal{P}_M es completo con esta norma.

Demostración. pendiente ■

A continuación un resultado esencial de aproximación, el cual se remite directamente a [8] para su demostración.

Teorema 1. El conjunto \mathcal{S} es denso en \mathcal{P}_M .

Lo anterior junto con la proposición 5.18 (referenciar) nos dice que

$$\mathbb{E} [(f \cdot M)_t^2(B)] \leq \|f\|_M^2, \quad \text{para } t \in (0, T], f \in \mathcal{S}, B \in \mathcal{B}(\mathbb{R}^d).$$

Luego, gracias a la cota anterior, si $(f_m)_{m \geq 1}$ es una sucesión de Cauchy en $(\mathcal{S}, \|\cdot\|_M)$, entonces la sucesión $((f_m \cdot M)_t(B))_{m \geq 1}$ es una sucesión de Cauchy en $L^2(\mathbb{P})$. Dado que ambos espacios son completos, se sigue la existencia de objetos en \mathcal{P} y $L^2(\mathbb{P})$, que denotamos por f y $(f \cdot M)_t(B)$ respectivamente, de tal forma que

$$f_m \xrightarrow[m \rightarrow \infty]{\|\cdot\|_M} f \implies (f_m \cdot M)_t(B) \xrightarrow[m \rightarrow \infty]{L^2(\mathbb{P})} (f \cdot M)_t(B).$$

De lo anterior, concluimos con el siguiente teorema.

Teorema 2. Sea M una medida martingala digna. Entonces para cualquier $f \in \mathcal{P}_M$, la integral $(f \cdot M)$ es una medida martingala digna que satisface la proposición 5.23 (referenciar). Más aún, para cualquier $t \in (0, T]$ y $A, B \in \mathcal{B}(\mathbb{R}^d)$,

$$\langle (f \cdot M)(A), (f \cdot M)(B) \rangle_t = \int_{A \times B \times (0, t]} f(x, s) f(y, s) Q_M(dx \, dy \, ds),$$

y además se tiene la siguiente cota en $L^2(\mathbb{P})$.

$$\mathbb{E} [(f \cdot M)_t^2(B)] \leq \|f\|_M^2.$$

De hecho, es posible extender la cota en $L^2(\mathbb{P})$ a una cota en $L^p(\mathbb{P})$, creando así una desigualdad del tipo Burkholder.

Teorema 3 (Desigualdades de Burkholder). Sea M una medida martingala digna. Entonces para cualquier $p \geq 2$ existe una constante $c_p \in (0, \infty)$ tal que para cualquier f predecible y cualquier $t > 0$,

$$\mathbb{E} [| (f \cdot M)_t(B) |^p] \leq c_p \mathbb{E} \left[\left(\int_{\mathbb{R}^d \times \mathbb{R}^d \times (0, T]} |f(x, t) f(y, t)| K_M(dx \, dy \, dt) \right)^{p/2} \right]$$

Con esto terminamos la construcción de la integral de Itô-Walsh. Ya nos es posible hablar de integrales con respecto a un ruido blanco \dot{W} , pues sabemos que este proceso en $\mathbb{R}^d \times [0, \infty)$ induce una medida martingala digna, y por lo tanto la noción de integral está bien definida para funciones aleatorias adecuadas.

1.3. Formulación Mild de una SPDE. Existencia y unicidad de las soluciones

Una vez que tenemos construida la teoría de la integral de Itô-Walsh, podemos pasar darle un significado riguroso a una ecuación diferencial parcial estocástica, así como definir con claridad qué significa resolverla. Como vimos en la introducción del capítulo, la heurística para llegar a un problema en cierto sentido equivalente a aquél que involucra derivadas sigue de cerca las ideas de las soluciones débiles en ecuaciones diferenciales parciales clásicas. Concretamente, el uso de integración por partes para lograr el objetivo es clave.

Consideremos la ecuación diferencial estocástica formal dada por

$$\begin{cases} \partial_t u(x, t) = \partial_{xx} u(x, t) + f(u(x, t)) \dot{W}, & \text{si } t > 0, x \in [0, L] \\ \partial_x u(0, t) = \partial_x u(L, t) = 0, & \text{si } t > 0, \\ u(x, 0) = u_0(x), & \text{si } x \in [0, L]. \end{cases}$$

donde \dot{W} es un ruido blanco con respecto a alguna filtración $(\mathcal{F}_t)_{t \geq 0}$, y $u_0 : [0, L] \rightarrow \mathbb{R}$ es una función determinista, medible y acotada, mientras que $f : \mathbb{R} \rightarrow \mathbb{R}$ es una función globalmente Lipschitz y acotada.

Si formalmente multiplicamos el problema por una función $\phi \in C^\infty(\mathbb{R})$ tal que $\phi(0) = \phi(L) = 0$ e integramos con respecto al tiempo y posteriormente con respecto al espacio, obtendremos que el problema anterior se puede reformular como

$$\begin{aligned} \int_0^L u(x, t) \phi(x) dx - \int_0^L u_0(x) \phi(x) dx &= \int_0^t \int_0^L \partial_{xx} u(x, s) \phi(x) dx ds \\ &+ \int_0^t \int_0^L f(u(x, s)) \phi(x) W(dx ds). \end{aligned}$$

Es destacable que ahora tenemos una ecuación integral estocástica en donde la integral con respecto al ruido blanco cobra total sentido. No obstante, lo anterior aún no puede ser tratado de manera rigurosa, ya que u al ser aleatorio no necesariamente será una función derivable. Por lo tanto, si de manera formal utilizamos integración por partes, al ser ϕ una función que se anula en la frontera, la integral de la parcial de u multiplicado por ϕ se transforma en

$$\int_0^t \int_0^L \partial_{xx} u(x, s) \phi(x) dx ds = \int_0^t \int_0^L u(x, s) \phi''(x) dx ds,$$

por lo que sustituyendo la expresión anterior en la reormulación de nuestro problema, obtenemos una ecuación integral estocástica. Tal ecuación integral tiene completo sentido, y dado que es una reformulación de nuestro problema inicial pero que puede tratarse de manera rigurosa, podemos considerar que una manera de resolver nuestro problema inicial, es resolver la reformulación integral anterior. Esta es conocida como la solución mild de nuestro problema anterior.

Definición 12. Decimos que u es una solución *mild* del problema inicial anterior, si existe un proceso estocástico $u(x, t)$ tal que para cualquier función $\phi \in C^\infty([0, L])$, donde $\phi'(0) = \phi'(L) = 0$, se tiene la igualdad

$$\int_0^L u(x, t) \phi(x) dx - \int_0^L u_0(x) \phi(x) dx = \int_0^t \int_0^L u(x, s) \phi''(x) dx ds + \int_0^t \int_0^L f(u(x, s)) \phi(x) W(dx ds).$$

Es conocida teoría que resuelve la ecuación diferencial estocástica del calor. En particular, se tiene el siguiente resultado.

Teorema 4. La ecuación estocástica del calor, con función f Lipschitz y acotada, tiene una única solución \mathbb{P} - casi seguramente, que satisface lo siguiente para cualquier $T > 0$.

$$\sup_{0 \leq x \leq L} \sup_{0 \leq t \leq T} \mathbb{E} [|u(x, t)|^2] < \infty.$$

Este es el resultado principal del capítulo. Y para su prueba, se requiere el siguiente lema bastante conocido en el ámbito de las ecuaciones diferenciales.

Teorema 5 (Lema de Gronwall). Supongamos $\phi_1, \phi_2, \dots : [0, T] \rightarrow [0, +\infty)$ son medibles y no decrecientes. Supongamos también que existe una constante $A \in \mathbb{R}$ tal que para cualquier entero $n \geq 1$ y para cualquier $t \in [0, T]$,

$$\phi_{n+1}(t) \leq A \int_0^t \phi_n(s) ds.$$

Entonces

$$\phi_n(t) \leq \phi_1(T) \frac{(At)^{n-1}}{(n-1)!}, \quad \text{para cualquier } n \geq 1, \text{ y } t \in [0, T].$$

Finalmente, el siguiente teorema nos muestra que no solamente hay existencia y unicidad de las solución a la ecuación estocástica del calor, sino que también existe una modificación continua de la misma.

Teorema 6. Sea $\tilde{u}(x, t)$ la solución única \mathbb{P} - casi seguramente de la ecuación estocástica del calor. Entonces existe una modificación $u(x, t)$ de dicha solución cuyas trayectorias son continuas.

Retomaremos la teoría desarrollada hasta aquí sobre ecuaciones diferenciales parciales estocásticas cuando lleguemos al capítulo 4, en donde estudiamos convergencia de las densidades de los promedios espaciales de la solución a la ecuación estocástica del calor.

Capítulo 2

Elementos de Cálculo de Malliavin

En este capítulo se revisa material esencial de cálculo de Malliavin. Esta disciplina nació a finales de la década de 1970 a partir de los trabajos de Paul Malliavin. La intención de Malliavin era hallar una prueba probabilista del teorema de Hörmander sobre el efecto regularizante de las soluciones de cierta ecuación diferencial estocástica, puesto que Lars Hörmander en 1967 encontró una prueba que utilizaba herramientas puramente analíticas. Con dicha motivación inicial, en las últimas décadas el cálculo de Malliavin se ha convertido en una poderosa herramienta matemática para estudiar problemas diversos en probabilidad.

En palabras de Nourdin y Peccati (ver prefacio de [7]), el cálculo de Malliavin se puede entender como un cálculo diferencial en infinitas dimensiones, cuyos operadores actúan sobre funcionales de procesos Gaussianos generales, y que en conjunto con el método de Stein, se obtiene una amalgama poderosa para deducir teoremas tipo límite central, así como tasas de convergencia explícitas para los mismos.

Tres operadores son esenciales en esta área: la derivada de Malliavin D , el operador de divergencia δ (o integral de Skorokhod) y el generador infinitesimal. El propósito de este capítulo es estudiar el material nuclear correspondiente a estos operadores.

2.1. Cálculo de Malliavin en el caso unidimensional

Siguiendo de cerca la exposición hecha por [7], comenzamos el estudio del tópico por el caso unidimensional. Aquí es más sencillo observar propiedades de estos objetos, los cuales se generalizan a dimensiones infinitas en la siguiente sección.

2.1.1. La derivada de Malliavin.

El operador con el que comenzamos nuestro estudio es la derivada de Malliavin. La idea es definir la derivada de variables aleatorias del tipo $f(N)$, donde f es una función determinista,

y N es una variable aleatoria con distribución Normal estándar.

En el contexto unidimensional, la definición de este operador se realiza utilizando ideas similares a aquellas que nos llevan a la definición de derivada débil en los espacios de Sobolev. Dichas ideas consisten, a grandes rasgos, en generalizar la noción de derivada a objetos más generales que las funciones diferenciables en el sentido usual. Para ello primero buscamos una clase de funciones diferenciables con propiedades nobles pero que a su vez sea suficientemente rica, y posteriormente extendemos el operador en algún sentido.

Dado que buscamos definir la derivada de funciones de variables normales estándar, es natural definir el operador derivada en un subconjunto de los espacios $L^q(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ con $q \in [1, \infty)$, pero dotados de la medida gaussiana estándar.

Dicho lo anterior, denotemos por $\gamma : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ a la medida mencionada, a saber,

$$\gamma(A) = \int_A \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Consideremos el espacio $L^q(\mathbb{R}, \mathcal{B}(\mathbb{R}), \gamma(dx))$, con $q \in [1, \infty)$, que durante esta subsección denotaremos simplemente como $L^q(\gamma)$. En este espacio, se valen algunas propiedades resultan de gran utilidad. Tal es el caso del siguiente lema.

Lema 1. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función absolutamente continua con respecto a la medida de Lebesgue y tal que $f' \in L^1(\gamma)$. Entonces la función $h : \mathbb{R} \rightarrow \mathbb{R}$ dada por $h(x) = xf(x)$ está en $L^1(\gamma)$ y se cumple que

$$\int_{\mathbb{R}} xf(x)d\gamma(x) = \int_{\mathbb{R}} f'(x)d\gamma(x).$$

Demostración. Supongamos primero que $f(0) = 0$. Probamos que $h \in L^1(\gamma)$. En efecto,

$$\begin{aligned} \int_{-\infty}^{\infty} |x||f(x)|d\gamma(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left| \int_0^x f'(y)dy \right| |x|e^{-\frac{x^2}{2}} dx \\ &\leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\int_0^x |f'(y)| dy \right) |x|e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \left(\int_x^0 |f'(y)| dy \right) (-x)e^{-\frac{x^2}{2}} dx \\ &\quad + \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \left(\int_0^x |f'(y)| dy \right) xe^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \int_{-\infty}^y |f'(y)| (-x)e^{-\frac{x^2}{2}} dx dy + \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \int_y^{\infty} |f'(y)| xe^{-\frac{x^2}{2}} dx dy \\ &= \int_{-\infty}^0 |f'(y)| \frac{1}{\sqrt{2\pi}} \left(e^{-\frac{x^2}{2}} \Big|_{-\infty}^y \right) dy + \int_0^{\infty} |f'(y)| \frac{1}{\sqrt{2\pi}} \left(-e^{-\frac{x^2}{2}} \Big|_y^{\infty} \right) dy \\ &= \int_{-\infty}^{\infty} |f'(y)|d\gamma(y) < \infty, \end{aligned}$$

en donde hemos hecho uso del teorema de Fubini en la tercera igualdad. Por lo tanto, la función $h(x) = xf(x)$ en efecto está en $L^1(\gamma)$. Realizando cuentas similares, obtenemos que

$$\begin{aligned}
\int_{-\infty}^{\infty} xf(x)d\gamma(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\int_0^x f'(y)dy \right) xe^{-\frac{x^2}{2}} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \left(\int_x^0 f'(y)dy \right) (-x)e^{-\frac{x^2}{2}} dx \\
&\quad + \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \left(\int_0^x f'(y)dy \right) xe^{-\frac{x^2}{2}} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \int_{-\infty}^y f'(y)(-x)e^{-\frac{x^2}{2}} dx dy + \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \int_y^{\infty} f'(y)xe^{-\frac{x^2}{2}} dx dy \\
&= \int_{-\infty}^0 f'(y) \frac{1}{\sqrt{2\pi}} \left(e^{-\frac{x^2}{2}} \Big|_{-\infty}^y \right) dy + \int_0^{\infty} f'(y) \frac{1}{\sqrt{2\pi}} \left(-e^{-\frac{x^2}{2}} \Big|_y^{\infty} \right) dy \\
&= \int_{-\infty}^{\infty} f(y)d\gamma(y),
\end{aligned}$$

como queríamos. Para el caso en el que $f(0) \neq 0$, notamos que ■

En la búsqueda de la generalización de la derivada a funciones más generales, definimos a la clase de Schwarz \mathcal{S} como el conjunto de funciones en $C^\infty(\mathbb{R}, \mathbb{R})$ tales que ella y sus derivadas tienen crecimiento a lo más polinomial, es decir,

$$\mathcal{S} := \left\{ f \in C^\infty(\mathbb{R}) : \forall p \in \{0, 1, 2, \dots\} \exists n \geq 1 \text{ tal que } \lim_{x \rightarrow \pm\infty} \frac{|f^{(p)}(x)|}{1 + |x|^n} < \infty \right\}.$$

Observamos que este conjunto dotado de la restricción correspondiente de la norma $L^q(\gamma)$ forma subespacio lineal de $L^q(\gamma)$, puesto que las combinaciones lineales de elementos infinitamente derivables vuelven a ser infinitamente derivables, y las combinaciones lineales también son compatibles con los límites. Es de destacar que este espacio es suficientemente amplio como para poseer a los polinomios. Más aún, este último conjunto es denso en $L^q(\gamma)$ para cualquier $q \in [1, \infty)$. Esto es consecuencia del siguiente lema.

Lema 2. Para cualquier $q \in [1, \infty)$, el conjunto de los monomios de la forma $\{x^n : n \geq 0\}$ genera un subespacio denso del espacio $L^q(\gamma)$. En particular, la clase de Schwarz \mathcal{S} forma un subconjunto denso de $L^q(\gamma)$.

Demostración. Un corolario del teorema de Hann-Banach caracteriza la propiedad de ser un subespacio lineal denso de un espacio L^q . En nuestro caso basta con probar que para cualquier elemento $g \in L^p(\gamma)$,

$$\int_{\mathbb{R}} g(x)x^n d\gamma(x) = 0, \quad \forall n \geq 0 \quad \implies \quad g = 0 \quad \gamma - \text{casi en todas partes},$$

donde $p \in (1, \infty]$ es el exponente conjugado de q . Sea pues $g \in L^p(\gamma)$ y supongamos que para cualquier $n \geq 0$, el antecedente de la implicación anterior es válido. Usaremos un argumento

que involucra la transformada de Fourier. Notamos primero que la transformada de $g(x)e^{-\frac{x^2}{2}}$ existe. En efecto, por la desigualdad de Hölder, para $t \in \mathbb{R}$ arbitrario,

$$\begin{aligned} \int_{\mathbb{R}} |g(x)e^{-\frac{x^2}{2}} e^{itx}| dx &\leq \int_{\mathbb{R}} |g(x)| e^{-\frac{x^2}{2}} dx \\ &= \int_{\mathbb{R}} |g(x)| d\gamma(x) \\ &\leq \|g\|_{L^p(\gamma)} \left(\int_{\mathbb{R}} d\gamma(x) \right)^{\frac{1}{q}} \\ &= \|g\|_{L^p(\gamma)} < \infty. \end{aligned}$$

Por otro lado, para cualquier $x \in \mathbb{R}$,

$$\left| g(x) \sum_{k=0}^n \frac{(itx)^k}{k!} \right| \leq |g(x)| \sum_{k=0}^n \frac{|tx|^k}{k!} \leq |g(x)| \sum_{k=0}^{\infty} \frac{|tx|^k}{k!} = |g(x)| e^{|tx|},$$

y la cota anterior vale para cualquier $n \geq 1$. Nuevamente usando la desigualdad de Hölder,

$$\int_{\mathbb{R}} \left| g(x) \sum_{k=0}^n \frac{(itx)^k}{k!} \right| d\gamma(x) \leq \int_{\mathbb{R}} |g(x)| e^{|tx|} d\gamma(x) \leq \|g\|_{L^p(\gamma)} \left(\int_{\mathbb{R}} (e^{|tx|})^q d\gamma(x) \right)^{\frac{1}{q}}.$$

Observamos que se tiene la igualdad

$$\int_{\mathbb{R}} (e^{|tx|})^q d\gamma(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{q|tx|} e^{-\frac{x^2}{2}} dx,$$

y notamos que esta última integral es finita para cualquier $q \in [1, \infty)$, ya que el término correspondiente a la medida gaussiana domina al otro término exponencial. Por lo tanto, $|g(x)|e^{|tx|}$ es una función en $L^1(\gamma)$ que domina a las sumas parciales anteriores. Así, usando el teorema de convergencia dominada,

$$\int_{\mathbb{R}} g(x) e^{-\frac{x^2}{2}} e^{itx} dx = \int_{\mathbb{R}} \lim_{n \rightarrow \infty} g(x) \sum_{k=0}^n \frac{(itx)^k}{k!} d\gamma(x) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{(it)^k}{k!} \int_{\mathbb{R}} g(x) x^k d\gamma(x) = 0,$$

donde hemos hecho uso de la hipótesis del enunciado en la última igualdad. Hemos hallado que la transformada de Fourier de $g(x)e^{-\frac{x^2}{2}}$ es idénticamente la función 0. Al ser la transformada de Fourier inyectiva en $L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), dx)$, se sigue que $g(x)e^{-\frac{x^2}{2}}$ es la función idénticamente 0 salvo conjuntos de medida de Lebesgue 0. En particular, g es la función idénticamente 0, γ — casi en todas partes.

Concluimos la prueba notando que los monomios son un subconjunto de la clase de Schwarz \mathcal{S} . ■

Sabiendo que \mathcal{S} es un subespacio lineal denso de $L^q(\gamma)$, podemos generalizar la derivada de las funciones en el espacio de Schwarz a funciones más generales que pertenecen a $L^q(\gamma)$. Para ello denotemos por ahora como $\frac{d^p}{dx^p}f$ a la derivada de orden p de una función $f \in \mathcal{S}$, para $p \in \{0, 1, 2, \dots\}$. Dicha derivada siempre está definida para tales funciones. Si pensamos en la derivación de orden p como aplicar un operador a un elemento del espacio de Schwarz, resulta que dicho operador es lineal. Dado que buscamos extender la derivada a funciones más generales, es natural pensar en la posibilidad de extender el operador $\frac{d^p}{dx^p} : \mathcal{S} \rightarrow L^q(\gamma)$ a un operador que sea cerrado en un dominio más grande que \mathcal{S} . Este es el contenido del siguiente teorema.

Teorema 7. Sea $\frac{d^p}{dx^p} : \mathcal{S} \rightarrow L^q(\gamma)$ el operador derivada de orden p . Dicho operador es cerrable para todo $q \in [1, \infty)$ y todo entero $p \geq 1$

Demostración. Usamos la siguiente caracterización de ser un operador cerrable. Supongamos que $(f_n)_{n \geq 1}$ es una sucesión en \mathcal{S} y que $g \in L^q(\gamma)$ son tales que

$$f_n \xrightarrow{n \rightarrow \infty} 0 \quad \text{y} \quad \frac{d^p}{dx^p} f_n \xrightarrow{n \rightarrow \infty} g.$$

Buscamos probar que $g = 0$, γ – casi en todas partes.

Supongamos primero el caso $q > 1$. Sea $h \in \mathcal{S}$. Definimos $\delta h(x) := xh(x) - h'(x)$ y para $n \geq 1$, $\delta^{n+1}h(x) := \delta^n h(x)$, donde $\delta^1 := \delta$. Observemos primero que para cualquier $n \in \{1, 2, \dots, p\}$, $\delta^n h \in \mathcal{S}$. Esto se sigue de que $xh(x) - h'(x)$ sigue siendo una función infinitamente diferenciable con crecimiento a lo más polinomial, y un argumento inductivo.

Ahora bien, usando la igualdad

$$\int_{\mathbb{R}} \left(\frac{d^{p-1}}{dx^{p-1}} f_n(x) h(x) \right)' d\gamma(x) = \int_{\mathbb{R}} \frac{d^p}{dx^p} f_n(x) h(x) d\gamma(x) + \int_{\mathbb{R}} \frac{d^{p-1}}{dx^{p-1}} f_n(x) h'(x) d\gamma(x),$$

la cual se obtiene con la derivada del producto, se tiene que

$$\begin{aligned} \int_{\mathbb{R}} h(x) g(x) d\gamma(x) &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \frac{d^p}{dx^p} f_n(x) h(x) d\gamma(x) \\ &= \lim_{n \rightarrow \infty} \left(\int_{\mathbb{R}} \left(\frac{d^{p-1}}{dx^{p-1}} f_n(x) h(x) \right)' d\gamma(x) - \int_{\mathbb{R}} \frac{d^{p-1}}{dx^{p-1}} f_n(x) h'(x) d\gamma(x) \right) \\ &= \lim_{n \rightarrow \infty} \left(\int_{\mathbb{R}} x \frac{d^{p-1}}{dx^{p-1}} f_n(x) h(x) d\gamma(x) - \int_{\mathbb{R}} \frac{d^{p-1}}{dx^{p-1}} f_n(x) h'(x) d\gamma(x) \right) \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \frac{d^{p-1}}{dx^{p-1}} f_n(x) (xh(x) - h'(x)) d\gamma(x) \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \frac{d^{p-1}}{dx^{p-1}} f_n(x) \delta h(x) d\gamma(x), \end{aligned}$$

donde en la primera igualdad hemos hecho uso del teorema de convergencia dominada, y en la tercera igualdad hemos hecho uso del lema 1. Utilizando un argumento inductivo y la

continuidad del valor absoluto, obtenemos que

$$\left| \int_{\mathbb{R}} h(x)g(x)d\gamma(x) \right| \leq \int_{\mathbb{R}} |h(x)g(x)| d\gamma(x) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} |f_n(x)\delta^p h(x)| d\gamma(x).$$

Por hipótesis, dado que la sucesión f_n converge a 0 en $L^q(\gamma)$ y $\delta^p h \in \mathcal{S} \subseteq L^{\frac{q}{q-1}}$ según el Lema 2, usando la desigualdad de Hölder,

$$\left| \int_{\mathbb{R}} h(x)g(x)d\gamma(x) \right| \leq \lim_{n \rightarrow \infty} \int_{\mathbb{R}} |f_n(x)\delta^p h(x)| d\gamma(x) \leq \lim_{n \rightarrow \infty} \|f_n\|_{L^q(\gamma)} \|\delta^p h\|_{L^{\frac{q}{q-1}}(\gamma)} = 0,$$

■

de donde deducimos que para cualquier $h \in \mathcal{S}$,

$$\int_{\mathbb{R}} h(x)g(x)d\gamma(x) = 0.$$

Dado que \mathcal{S} es denso en $L^q(\gamma)$, se sigue que $g = 0$, γ – casi en todas partes.

Hemos demostrado que el operador derivada de orden p es cerrable en \mathcal{S} . Por lo tanto, por definición existe un operador cerrado, que vamos a denotar por D^p , con dominio $\text{Dom}(D^p)$ un subespacio lineal de $L^q(\gamma)$ tal que $\mathcal{S} \subseteq \text{Dom}(D^p)$ y $D^p f = \frac{d^p}{dx^p} f$ para cualquier $f \in \mathcal{S}$. Estos elementos nos permiten definir la derivada de Malliavin.

Definición 13. Sea $p \geq 1$ un entero positivo. Definimos la derivada de Malliavin como el operador D^p que extiende a la derivada $\frac{d^p}{dx^p}$ vista como operador de \mathcal{S} a $L^q(\gamma)$, para $q \in [1, \infty)$.

Denotaremos al dominio de D^p como $\mathbb{D}^{p,q} := \text{Dom}(D^p)$. A dichos espacios también se les conoce como espacios de Watanabe-Sobolev. Es conocido que al ser D^p un operador cerrado, podemos definir la *norma de la gráfica* de dicho operador en su dominio como la función $\|\cdot\|_G : \mathbb{D}^{p,q} \rightarrow [0, \infty)$ dada por,

$$\|f\|_G := \|f\|_{L^q(\gamma)} + \|D^p f\|_{L^q(\gamma)}.$$

Dicha norma vuelve a la derivada de Malliavin D^p un operador continuo. Más aún, es un resultado conocido de análisis funcional que, al ser D^p un operador cerrado, su dominio $\mathbb{D}^{p,q}$ es un espacio de Banach con la norma $\|\cdot\|_G$ definida antes.

Ahora bien, es natural pensar en lo que sucede con una aplicación sucesiva de las derivadas de Malliavin. Se tiene que para cualquier $p \geq 1$ entero, y $f \in \mathbb{D}^{p+1,q}$,

$$D^{p+1}f = D(D^p f) = D^p(Df),$$

lo que nos dice que la derivada de Malliavin en efecto respeta la composición, tal y como ocurre en la derivación usual. Se sigue de esto que, al ser consistente la derivación sucesiva en el sentido de Malliavin, para una función $f \in \mathbb{D}^{p,q}$, las derivadas $D^k f$ existen para cualquier

$k \in \{1, \dots, p\}$, y por lo tanto, las normas $\|D^k f\|_{L^2(\gamma)}$ son finitas para cualquier $k \in \{1, \dots, p\}$. Luego, se puede demostrar que la función $\|\cdot\|_{\mathbb{D}^{p,q}} : \mathbb{D}^{p,q} \longrightarrow [0, \infty)$ dada por

$$\|f\|_{\mathbb{D}^{p,q}} := \left(\|f\|_{L^q(\gamma)}^q + \sum_{k=1}^p \|D^k f\|_{L^q(\gamma)}^q \right)^{\frac{1}{q}}$$

es una norma equivalente a la norma de la gráfica definida en $\mathbb{D}^{p,q}$ antes. Por lo tanto, esta norma también hace a $\mathbb{D}^{p,q}$ un espacio de Banach.

Concluimos observando que $\mathbb{D}^{p,q}$ también puede ser obtenido como la cerradura del subespacio \mathcal{S} con respecto a la norma $\|\cdot\|_{\mathbb{D}^{p,q}}$ definida antes, pero restringida al espacio \mathcal{S} , en donde la derivada de Malliavin coincide con la derivada usual.

Lo anterior permite hacer una caracterización del espacio $\mathbb{D}^{p,q}$ y una definición alternativa de la derivada de Malliavin: una función f está en dicho espacio si y solo si existe una sucesión de funciones $(f_n)_{n \geq 1} \subseteq \mathcal{S}$ tales que $f_n \xrightarrow[n \rightarrow \infty]{} f$ y $(\frac{d^p}{dx^p} f_n)_{n \geq 1}$ es una sucesión de Cauchy en $L^q(\gamma)$, para cualquier $k \in \{1, \dots, p\}$. Para tal sucesión de funciones, definimos

$$D^k f := \lim_{n \rightarrow \infty} \frac{d^k}{dx^k} f_n,$$

en donde el límite es en sentido $L^q(\gamma)$, y dicha función límite coincide con la derivada de Malliavin definida antes.

Este operador es de suma importancia. En la siguiente sección veremos que en el caso en que $q = 2$, la derivada de Malliavin vista como operador posee propiedades interesantes.

2.1.2. El operador de divergencia

A lo largo de esta sección consideraremos el espacio $L^2(\gamma)$, el cual es un espacio de Hilbert con producto interno dado por

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x)g(x) d\gamma(x).$$

A partir de ahora, y para el resto del texto, denotaremos simplemente por D^p a la derivada de Malliavin de orden p , aún cuando nos estemos refiriendo a la derivada de orden p convencional de una función f en el espacio de Schwarz.

Hasta ahora tenemos que la derivada de Malliavin es un operador $D^p : \mathbb{D}^{p,2} \subseteq L^2(\gamma) \longrightarrow L^2(\gamma)$ y dicho operador, es cerrado por construcción, extiende al operador derivada convencional D^p definido en $\mathcal{S} \subseteq L^2(\gamma)$.

Siendo $L^2(\gamma)$ un espacio de Hilbert, si denotamos por $\text{Dom}(\delta^p)$ como el conjunto de funciones $g \in L^2(\gamma)$ tales que existe una constante $c_g > 0$ y se cumple que

$$|\langle D^p f, g \rangle_{L^2(\gamma)}| \leq c_g \|f\|_{L^2(\gamma)}, \quad \forall f \in \mathbb{D}^{p,2}, x d$$

entonces para cada $g \in \text{Dom}(\delta^p)$ el funcional lineal, con dominio $\mathbb{D}^{2,q}$ y con regla de correspondencia $f \mapsto \langle D^p f, g \rangle$, es acotado precisamente por la condición anterior. De esta forma, por el teorema de Hanh-Banach podemos extender dicho funcional a un funcional lineal acotado cuyo dominio es todo $L^2(\gamma)$.

Con el funcional anterior podemos usar el teorema de representación de Riesz para obtener la existencia de un único elemento en $L^2(\gamma)$, que denotaremos por $\delta^p g$ tal que

$$\langle D^p f, g \rangle_{L^2(\gamma)} = \langle f, \delta^p g \rangle_{L^2(\gamma)}, \quad \forall f \in L^2(\gamma).$$

En términos de integrales, se tiene que para cualquier $g \in \text{Dom}(\delta^p)$, existe un único elemento $\delta^p g \in L^2(\gamma)$ tal que

$$\int_{\mathbb{R}} D^p f(x) g(x) d\gamma(x) = \int_{\mathbb{R}} f(x) \delta^p g(x) d\gamma(x), \quad \forall f \in L^2(\gamma).$$

Con los elementos anteriores, podemos definir el operador de divergencia.

Definición 14. Sea $p \geq 1$ un entero positivo. Denotamos

$$\text{Dom}(\delta^p) := \{g \in L^2(\gamma) : |\langle D^p f, g \rangle_{L^2(\gamma)}| \leq c \|f\|_{L^2(\gamma)}, \quad \forall f \in \mathbb{D}^{2,p}\},$$

y definimos el operador de divergencia p -ésimo como la aplicación $\delta^p : \text{Dom}(\delta^p) \rightarrow L^2(\gamma)$ tal que asigna a cada $g \in \text{Dom}(\delta^p)$ al único elemento $\delta^p g \in L^2(\gamma)$ que, gracias al teorema de representación de Riesz, cumple la igualdad

$$\int_{\mathbb{R}} D^p f(x) g(x) d\gamma(x) = \int_{\mathbb{R}} f(x) \delta^p g(x) d\gamma(x), \quad \forall f \in L^2(\gamma).$$

En el caso en que $p = 1$, directamente escribimos $\delta^1 = \delta$. Observamos que el operador δ^p es directamente el operador adjunto de D^p , y como tal, es un operador cerrado.

Es de interés encontrar una forma de calcular el operador δ aplicado a alguna función en $\text{Dom}(\delta)$. En este sentido, existe una fórmula para calcular δg para $g \in \mathbb{D}^{1,2}$. Este es el contenido del siguiente lema.

Lema 3. Se cumple que $\mathcal{S} \subseteq \text{Dom}(\delta)$ y para cualquier $g \in \mathbb{D}^{1,2}$,

$$\delta g(x) = xg(x) - Dg(x) = xg(x) - g'(x).$$

En particular, para cualquier $g \in \mathcal{S}$, se cumple la fórmula anterior.

Demostración. Nótese que para $f, g \in \mathcal{S}$, por el lema 1,

$$\int_{\mathbb{R}} (f(x)g(x))' d\gamma(x) = \int_{\mathbb{R}} x f(x) g(x) d\gamma(x),$$

por lo que

$$\int_{\mathbb{R}} xf(x)g(x)d\gamma(x) = \int_{\mathbb{R}} D(f(x)g(x))d\gamma(x) = \int_{\mathbb{R}} Df(x)g(x)d\gamma(x) + \int_{\mathbb{R}} f(x)Dg(x)d\gamma(x),$$

así que restando,

$$\int_{\mathbb{R}} Df(x)g(x)d\gamma(x) = \int_{\mathbb{R}} xf(x)g(x)d\gamma(x) - \int_{\mathbb{R}} f(x)Dg(x)d\gamma(x),$$

y por definición de δg , se tiene que

$$\int_{\mathbb{R}} f(x)\delta g(x)d\gamma(x) = \int_{\mathbb{R}} xf(x)g(x)d\gamma(x) - \int_{\mathbb{R}} f(x)Dg(x)d\gamma(x).$$

Pero lo anterior ocurre para cualquier $f \in \mathcal{S}$. Esto nos dice que $\mathcal{S} \subseteq \text{Dom}(\delta)$ y que para cualquier $g \in \mathcal{S}$, la fórmula anterior es válida.

Utilizando un argumento de aproximación, tenemos que para cualquier $g \in \mathbb{D}^{1,2}$, $\delta g(x) = xg(x) - Dg(x)$. ■

2.1.3. El semigrupo de Ornstein-Uhlenbeck

El semigrupo de Ornstein-Uhlenbeck es una generalización del operador Laplaciano a dimensión infinita. Comenzamos directamente definiendo el semigrupo de Ornstein-Uhlenbeck.

Definición 15. El semigrupo de Ornstein-Uhlenbeck, denotado por $(P_t)_{t \geq 0}$, se define como una familia de operadores del espacio \mathcal{S} dotado de la norma $\|\cdot\|_{L^q(\gamma)}$ a $L^q(\gamma)$ como sigue: para $f \in \mathcal{S}$ y $t \geq 0$,

$$P_t f(x) = \int_{\mathbb{R}} f\left(e^{-t}x + \sqrt{1 - e^{-2t}}y\right) d\gamma(y), \quad \forall x \in \mathbb{R}.$$

De manera inmediata se observa que esta es una familia de operadores lineales. Tal como se puede esperar, esta familia de operadores cumple la propiedad de semigrupo, y está directamente relacionada con el proceso de Ornstein-Uhlenbeck. Ambos hechos se exploran más adelante. Comenzamos estudiando propiedades básicas de $(P_t)_{t \geq 0}$.

Teorema 8. Sea $f \in \mathcal{S}$ y $q \in [1, \infty)$. Entonces se cumplen las siguientes proposiciones:

- $P_0 f(x) = f(x)$.
- $P_\infty f(x) := \lim_{t \rightarrow \infty} P_t f(x) = \int_{\mathbb{R}} f(y) d\gamma(y)$.
- $\int_{\mathbb{R}} |P_t f(x)|^q d\gamma(x) \leq \int_{\mathbb{R}} |f(x)|^q d\gamma(x)$.

Demostración. ■ Calculamos el semigrupo en $t = 0$. Dado que $\gamma(\mathbb{R}) = 1$, se tiene que

$$P_0 f(x) = \int_{\mathbb{R}} f\left(e^0 x + \sqrt{1 - e^0} y\right) d\gamma(y) = \int_{\mathbb{R}} f(x) d\gamma(y) = f(x).$$

- Como $f \in \mathcal{S}$, tiene crecimiento a lo más polinomial. Por lo tanto existe un real $M > 0$ y un natural $n \geq 1$ tal que para cualquier $y \in \mathbb{R} \setminus [-M, M]$, $|f(y)| \leq C(1 + |y|^n)$, para alguna constante $C > 0$. Observamos que esta última cota es una función en $L^q(\gamma)$. Por otro lado, es claro que en $[-M, M]$, la función $|f|$ alcanza su máximo, por lo que $\|f \mathbb{1}_{[-M, M]}\|_{\infty} < \infty$.

Finalmente, notamos que la función $g(y) = e^{-t}x + \sqrt{1 - e^{-2t}}y$ es una función continua y por lo tanto medible, así que

$$\begin{aligned} \left| f\left(e^{-t}x + \sqrt{1 - e^{-2t}}y\right) \right| &\leq C(1 + |y|^n) \mathbb{1}_{g^{-1}[-M, M]^c}(y) + \|f \mathbb{1}_{[-M, M]}\|_{\infty} \mathbb{1}_{g^{-1}[-M, M]}(y) \\ &\leq C(1 + |y|^n) + \|f \mathbb{1}_{[-M, M]}\|_{\infty}, \end{aligned}$$

función que está en $L^1(\gamma)$, y que es independiente de t . Por lo tanto, por el teorema de convergencia dominada,

$$P_{\infty} f(x) = \lim_{t \rightarrow \infty} P_t f(x) = \int_{\mathbb{R}} \lim_{t \rightarrow \infty} f\left(e^{-t}x + \sqrt{1 - e^{-2t}}y\right) d\gamma(y) = \int_{\mathbb{R}} f(y) d\gamma(y).$$

- Para $q \in [1, \infty)$, por definición,

$$\begin{aligned} \int_{\mathbb{R}} |P_t f(x)|^q d\gamma(x) &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} f\left(e^{-t}x + \sqrt{1 - e^{-2t}}y\right) d\gamma(y) \right|^q d\gamma(x) \\ &\leq \int_{\mathbb{R}} \int_{\mathbb{R}} \left| f\left(e^{-t}x + \sqrt{1 - e^{-2t}}y\right) \right|^q d\gamma(y) d\gamma(x) \\ &= \mathbb{E} \left[f \left(\left| e^{-t}X + \sqrt{1 - e^{-2t}}Y \right|^q \right) \right], \end{aligned}$$

donde hemos hecho uso de la desigualdad de Jensen, y X, Y son variables aleatorias normales estándar independientes. Pero notemos que, denotando $Z := e^{-t}X + \sqrt{1 - e^{-2t}}Y$, la función característica de Z es

$$\phi_Z(z) = \mathbb{E} \left[e^{ize^{-t}X + iz\sqrt{1 - e^{-2t}}Y} \right] = \mathbb{E} \left[e^{ize^{-t}X} \right] \mathbb{E} \left[e^{iz\sqrt{1 - e^{-2t}}Y} \right] = e^{-\frac{z^2 e^{-2t}}{2}} e^{-\frac{z^2 (1 - e^{-2t})}{2}} = e^{-\frac{z^2}{2}},$$

por lo que Z tiene una distribución normal estándar. Se sigue que

$$\int_{\mathbb{R}} |P_t f(x)|^q d\gamma(x) = \mathbb{E} \left[f \left(\left| e^{-t}X + \sqrt{1 - e^{-2t}}Y \right|^q \right) \right] \leq \mathbb{E} [|f(Z)|^q] = \int_{\mathbb{R}} |f(x)|^q d\gamma(x).$$

■

Del último ítem de la proposición anterior se deduce que para cualquier $f \in \mathcal{S}$ y cualquier $t \geq 0$, $P_t f \in L^q(\gamma)$. Más aún, la familia de operadores $(P_t)_{t \geq 0}$ que van del espacio $(\mathcal{S}, \|\cdot\|_{L^q(\gamma)})$ a $L^q(\gamma)$ son uniformemente acotados por la constante 1.

Al ser \mathcal{S} un subconjunto denso de $L^q(\gamma)$ y ser este último un espacio completo, por el conocido teorema BLT de análisis funcional, los operadores de $(P_t)_{t \geq 0}$ se pueden extender de manera única a una familia de operadores, que también denotaremos por $(P_t)_{t \geq 0}$, con dominio todo el espacio $L^q(\gamma)$, y cuyas normas siguen siendo acotadas por la constante 1. Esto último hace a $(P_t)_{t \geq 0}$ una familia de operadores lineales contractivos.

Probamos ahora que $(P_t)_{t \geq 0}$ en efecto forma un semigrupo.

Teorema 9. Sean $s, t \geq 0$. Se tiene que $P_t P_s = P_{t+s}$ en $L^q(\gamma)$, para $q \in [1, \infty)$.

Demostración. Dado que la medida del espacio es finita, basta con probar el resultado para $L^1(\gamma)$. Sea $f \in L^1(\gamma)$. Tenemos las siguientes igualdades para $x \in \mathbb{R}$:

$$\begin{aligned} P_t(P_s f)(x) &= \int_{\mathbb{R}} P_s f \left(e^{-t}x + \sqrt{1 - e^{-2t}}y \right) d\gamma(x) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f \left(e^{-s}(e^{-t}x + \sqrt{1 - e^{-2t}}y) + \sqrt{1 - e^{-2s}}z \right) d\gamma(z) d\gamma(y) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f \left(e^{-(t+s)}x + e^{-s}\sqrt{1 - e^{-2t}}y + \sqrt{1 - e^{-2s}}z \right) d\gamma(z) d\gamma(y) \\ &= \mathbb{E} \left[f \left(e^{-(t+s)}x + e^{-s}\sqrt{1 - e^{-2t}}Y + \sqrt{1 - e^{-2s}}Z \right) \right], \end{aligned}$$

donde Y y Z denotan variables aleatorias normales estándar e independientes. Ahora bien, nótese que la función característica de la variable $W := e^{-s}\sqrt{1 - e^{-2t}}Y + \sqrt{1 - e^{-2s}}Z$ está dada por

$$\begin{aligned} \phi_W(w) &= \mathbb{E} \left[\exp \left\{ iw(e^{-s}\sqrt{1 - e^{-2t}}Y + \sqrt{1 - e^{-2s}}Z) \right\} \right] \\ &= \mathbb{E} \left[\exp \left\{ iwe^{-s}\sqrt{1 - e^{-2t}}Y \right\} \right] \mathbb{E} \left[\exp \left\{ iw\sqrt{1 - e^{-2s}}Z \right\} \right] \\ &= \exp \left\{ \frac{-w^2 e^{-2s}(1 - e^{-2t})}{2} \right\} \exp \left\{ \frac{-w^2(1 - e^{-2s})}{2} \right\} \\ &= \exp \left\{ \frac{-w^2(e^{-2s} - e^{-2(t+s)} + 1 - e^{-2s})}{2} \right\} \\ &= \exp \left\{ \frac{-w^2(1 - e^{-2(t+s)})}{2} \right\}, \end{aligned}$$

la cual corresponde con la función característica de una variable $V := \sqrt{1 - e^{-2(t+s)}}N$, donde

N es gaussiana estándar. Se sigue que

$$\begin{aligned}
P_t(P_s f)(x) &= \mathbb{E} \left[f \left(e^{-(t+s)} x + e^{-s} \sqrt{1 - e^{-2t}} Y + \sqrt{1 - e^{-2s}} Z \right) \right] \\
&= \mathbb{E} \left[f \left(e^{-(t+s)} + \sqrt{1 - e^{-2(t+s)}} V \right) \right] \\
&= \int_{\mathbb{R}} f \left(e^{-(t+s)} + \sqrt{1 - e^{-2(t+s)}} v \right) d\gamma(v) \\
&= P_{t+s} f(x).
\end{aligned}$$

■

Finalmente, la siguiente proposición nos dice que tanto P_t como D , vistos como operadores, pueden ser intercambiados en el espacio $\mathbb{D}^{1,2}$.

Teorema 10. Sea $f \in \mathbb{D}^{1,2}$ y $t \geq 0$. Entonces $P_t f \in \mathbb{D}^{1,2}$ y $DP_t f = e^{-t} P_t Df$.

Demostración. Sea $f \in \mathcal{S}$. La derivada de Malliavin coincide con la derivada usual en dicho espacio y por lo tanto, dado que la función $e^{-t} f' (e^{-t} x + \sqrt{1 - e^{-2t}} y)$ está dominada por una función similar a la del ítem 3 del teorema 8, podemos derivar bajo el signo de la integral, hallando que

$$\begin{aligned}
DP_t f(x) &= \frac{d}{dx} \int_{\mathbb{R}} f \left(e^{-t} x + \sqrt{1 - e^{-2t}} y \right) d\gamma(y) \\
&= \int_{\mathbb{R}} e^{-t} \frac{d}{dx} f \left(e^{-t} x + \sqrt{1 - e^{-2t}} y \right) d\gamma(y) \\
&= e^{-t} P_t Df(x),
\end{aligned}$$

para cualquier $x \in \mathbb{R}$. Esto prueba la propiedad en $\mathcal{S} \subseteq \mathbb{D}^{1,2}$. Para f en dicho espacio general, usamos que \mathcal{S} es denso en $\mathbb{D}^{1,2}$. ■

Antes de continuar con la siguiente sección, mostramos la conexión de esta familia de operadores con el semigrupo de Ornstein-Uhlenbeck:

Teorema 11. Sea B un movimiento browniano estándar. Consideramos la ecuación diferencial estocástica dada por

$$dX_t^x = \sqrt{2} dB_t - X_t^x dt, \quad t \geq 0,$$

donde $X_0^x = x \in \mathbb{R}$. Se tiene que la solución de la ecuación diferencial anterior está dada por

$$X_t^x = e^{-t} x + \sqrt{2} \int_0^t e^{-(t-s)} dB_s, \quad t \geq 0,$$

el cual es un proceso de Markov. Más aún, el semigrupo asociado a dicho proceso es justamente $(P_t)_{t \geq 0}$. Tal proceso es justamente el proceso de Ornstein-Uhlenbeck.

Demostración. Comenzamos notando que en efecto el proceso anteriormente definido cumple que $X_0^x = x$. Pero esto es claro ya que

$$X_0^x = e^0 x + \int_0^0 e^{-(t-s)} dB_s = x.$$

Vemos ahora que se cumple la ecuación diferencial estocástica. Sea $t > 0$ y nótese que

$$\begin{aligned} X_t^x &= e^{-t} x + \sqrt{2} \int_0^t e^{-(t-s)} dB_s \\ &= e^{-t} x + \sqrt{2} e^{-t} \int_0^t e^s dB_s \\ &= e^{-t} \left(X_0^x e^0 + \sqrt{2} e^{-t} \int_0^t e^s dB_s \right), \end{aligned}$$

de donde concluimos que

$$e^t X_t^x - e^0 X_0^x = \sqrt{2} \int_0^t e^s dB_s,$$

lo cual en notación diferencial equivale a que

$$d(e^t X_t^x) = \sqrt{2} e^t dB_t, \quad t > 0.$$

Usando ahora la regla del producto y el hecho de que e^{-t} es un proceso de variación finita,

$$\begin{aligned} dX_t^x &= d(e^{-t} e^t X_t^x) \\ &= e^{-t} d(e^t X_t^x) + e^t X_t^x d(e^{-t}) + 0 \\ &= e^{-t} \sqrt{2} e^t dB_t - e^t X_t^x e^{-t} dt \\ &= \sqrt{2} dB_t - X_t^x dt, \end{aligned}$$

tal como queríamos probar. Por lo tanto el proceso $(X_t^x)_{t \geq 0}$ definido como antes es solución a la ecuación diferencial estocástica dada. Como tal, dicho proceso resulta ser una difusión, y en particular un proceso de Markov. Resta ver que su semigrupo justamente es el semigrupo de Ornstein-Uhlenbeck definido antes.

Directamente de la expresión de X_t^x , observamos que tenemos una integral de Itô cuyo integrando es estocástico. Por lo tanto, dicha integral forma un proceso Gaussiano adaptado a la filtración generada por B . Por otro lado, calculando la función de medias y covarianzas del proceso hallamos lo siguiente:

$$\blacksquare \mu(t) = \mathbb{E}[X_t^x] = e^{-t} x$$

■

$$\begin{aligned}
\Gamma(s, t) &= \text{Cov}(X_s^x, X_t^x) \\
&= \mathbb{E}[(X_t^x - e^{-t}x)(X_s^x - e^{-s}x)] \\
&= \mathbb{E}\left[\left(\sqrt{2}e^{-t} \int_0^t e^u dB_u\right)\left(\sqrt{2}e^{-s} \int_0^s e^v dB_v\right)\right] \\
&= 2e^{-(t+s)} \mathbb{E}\left[\int_0^{s \wedge t} e^{2u} du\right] \\
&= \frac{2}{2} e^{-(t+s)} (e^{2s \wedge t} - 1) \\
&= e^{-|t-s|} - e^{-(t+s)}.
\end{aligned}$$

En particular, se tiene para $s = t$ que $X_t^x \sim \text{Normal}(e^{-t}x, 1 - e^{-2t})$. Finalmente, si $B = (B_t)_{t \geq 0}$ representa el movimiento browniano sobre el cual $(X_t^x)_{t \geq 0}$ está definido, notamos que para $t \geq 0$,

$$\mathbb{E}[e^{-t}x + e^{-t}B_{e^{2t}-1}] = e^{-t}x$$

y además,

$$\mathbb{E}[(e^{-t}B_{e^{2t}-1})(e^{-s}B_{e^{2s}-1})] = e^{-(t+s)}(\min(e^{2t}, e^{2s}) - 1) = e^{-(t+s)}(e^{2s \wedge t} - 1) = e^{-|t-s|} - e^{-(t+s)}.$$

Concluimos que el proceso de Ornstein-Uhlenbeck $(X_t^x)_{t \geq 0}$ como solución a la ecuación diferencial estocástica dada antes, puede verse como

$$X_t^x = e^{-t}x - e^{-t}B_{e^{2t}-1}, \quad t \geq 0.$$

De lo anterior, es muy sencillo ver que para f medible y acotada,

$$\begin{aligned}
P_t f(x) &= \mathbb{E}[f(X_{t+s}^x) | \mathcal{F}_s] \\
&= \mathbb{E}_x[f(e^{-(t+s)}x - e^{-(t+s)}B_{e^{2(t+s)}-1}) | \mathcal{F}_s] \\
&= \int_{\mathbb{R}} f(e^{-t}x - \sqrt{1 - e^{-2t}}y) d\gamma(y),
\end{aligned}$$

el cual es justamente el semigrupo que hemos estado trabajando. ■

2.1.4. El generador infinitesimal

Como es clásico en el contexto de procesos de Markov, asociado a un semigrupo tenemos la noción de generador infinitesimal. En nuestro caso consideremos al conjunto siguiente:

$$\text{Dom}(L) := \left\{ f \in L^2(\gamma) : \lim_{h \rightarrow 0} \frac{P_h f - f}{h} \text{ existe en } L^2(\gamma) \right\}.$$

Dicho es directamente el dominio del generador infinitesimal L . Así, dada $f \in \text{Dom}(L)$, se tiene que

$$Lf := \lim_{h \rightarrow 0} \frac{P_h f - f}{h}.$$

Ya que el límite anterior recrea la definición de derivada en funciones reales, denotaremos

$$\frac{d}{dt} P_t f := \lim_{h \rightarrow 0} \frac{P_{t+h} f - P_t f}{h}, \quad \text{y en particular} \quad Lf := \left. \frac{d}{dt} P_t f \right|_{t=0},$$

omitiendo en ocasiones la función f cuando estemos hablando de una función en un conjunto particular. Ahora bien, en el espacio \mathcal{S} , para cualquier $t \geq 0$ se tiene el siguiente resultado:

$$\frac{d}{dt} P_t = \lim_{h \rightarrow 0} \frac{P_{t+h} - P_t}{h} = \lim_{h \rightarrow 0} P_t \frac{P_h - Id}{h} = P_t \lim_{h \rightarrow 0} \frac{P_h - Id}{h} = P_t \left. \frac{d}{dh} P_h \right|_{h=0} = P_t L.$$

Invirtiendo la manera en cómo se toman los límites, se obtiene también que

$$\frac{d}{dt} P_t = L P_t,$$

así que para cualquier $t \geq 0$, y $f \in \mathcal{S}$, tenemos que

$$\frac{d}{dt} P_t = P_t L = L P_t.$$

Finalmente, una proposición sumamente importante es el siguiente teorema.

Teorema 12. En el espacio de Schwarz, se tiene que $Lf = -\delta Df$. Más específicamente, para cualquier $f \in \mathcal{S}$,

$$Lf(x) = -x f'(x) + f''(x).$$

Demostración. Sea $f \in \mathcal{S}$ y $t > 0$. Notamos que, utilizando una función dominante, que existe gracias al hecho de que f es suave, intercambiamos derivada con integral, por lo que

$$\begin{aligned} \frac{d}{dt} P_t f(x) &= \frac{d}{dt} \int_{\mathbb{R}} f(e^{-t}x + \sqrt{1 - e^{-2t}}y) d\gamma(y) \\ &= \int_{\mathbb{R}} f'(e^{-t}x + \sqrt{1 - e^{-2t}}y) \left(-e^{-t}x + \frac{-y(-2e^{-2t})}{2\sqrt{1 - e^{-2t}}} \right) d\gamma(y) \\ &= \int_{\mathbb{R}} f'(e^{-t}x + \sqrt{1 - e^{-2t}}y) \left(-e^{-t}x + \frac{ye^{-2t}}{\sqrt{1 - e^{-2t}}} \right) d\gamma(y) \\ &= -x \int_{\mathbb{R}} f'(e^{-t}x + \sqrt{1 - e^{-2t}}y) e^{-t} d\gamma(y) + \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} \int_{\mathbb{R}} y f'(e^{-t}x + \sqrt{1 - e^{-2t}}y) d\gamma(y). \end{aligned}$$

Reconocemos en el primer integrando la derivada con respecto a x de $P_t f$, mientras que en el segundo término reconocemos la integral de una función de la forma $yg(y)$, donde $g(y) =$

$f'(e^{-t}x + \sqrt{1 - e^{-2t}}y)$, función que está en $L^1(\gamma)$. Por lo tanto, por el Lema 1,

$$\begin{aligned}
& -x \int_{\mathbb{R}} f'(e^{-t}x + \sqrt{1 - e^{-2t}}y) e^{-t} d\gamma(y) + \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} \int_{\mathbb{R}} y f'(e^{-t}x + \sqrt{1 - e^{-2t}}y) d\gamma(y) \\
& = -x \int_{\mathbb{R}} \frac{d}{dx} f(e^{-t}x + \sqrt{1 - e^{-2t}}y) d\gamma(y) + \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} \int_{\mathbb{R}} \frac{d}{dy} f(e^{-t}x + \sqrt{1 - e^{-2t}}y) d\gamma(y) \\
& = -x \frac{d}{dx} \int_{\mathbb{R}} f(e^{-t}x + \sqrt{1 - e^{-2t}}y) d\gamma(y) + \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} \sqrt{1 - e^{-2t}} \int_{\mathbb{R}} f''(e^{-t}x + \sqrt{1 - e^{-2t}}y) d\gamma(y) \\
& = -x(P_t f(x))' + e^{-2t} P_t f''(x).
\end{aligned}$$

Tomando límite cuando $t \rightarrow 0$ hallamos que

$$Lf(x) = \left. \frac{d}{dt} P_t f(x) \right|_{t=0} = -(xf'(x) + f''(x)).$$

Finalmente, dado que $f \in \mathcal{S}$, su derivada de Malliavin coincide con su derivada usual, por lo que usando la proposición ???, el generador infinitesimal de una función en el espacio de Schwarz viene dado por

$$Lf(x) = -\delta Df(x).$$

■

Como ejemplo de la potencia de la proposición anterior, se demuestra a continuación la desigualdad de Poincaré en el contexto de los espacios de Watanabe-Sobolev. Para tener una idea de la comparación entre ambos contextos, presentamos tanto la desigualdad de Poincaré en el contexto de espacios de Sobolev clásicos en \mathbb{R} como la desigualdad de Poincaré en $\mathbb{D}^{2,p}$.

Recordemos que $W^{k,p}(U) = \{f \in L^p(U, \mathcal{B}(U), dx) : D^j f \in L^p(U, \mathcal{B}(U), dx), j \in \{1, 2, \dots, k\}\}$ es el espacio de Sobolev que consiste en el conjunto de aquellas funciones en $L^p(U)$ ($U \subseteq \mathbb{R}$ subconjunto abierto) tales que todas sus derivadas hasta la k -ésima en el sentido distribucional también pertenecen a $L^p(U, \mathcal{B}(U), dx)$.

Teorema 13. (Desigualdad de Poincaré clásica) Sean $p \in [1, \infty)$ y $W_0^{1,p}(U) = \overline{C_c^\infty(U)}$ la cerradura del espacio de las funciones test, con respecto a la norma $\|\cdot\|_{W^{1,p}(U)}$ definida como

$$\|f\|_{W^{1,p}(U)} = \left(\|f\|_{L^p(U)}^p + \|Df\|_{L^p(U)}^p \right)^{1/p}, \quad \text{para } f \in W^{1,p}(U).$$

Entonces si U es un abierto acotado de \mathbb{R} , y $u \in W_0^{1,p}(U)$, se tiene la desigualdad

$$\|u\|_{L^p(U)} \leq C \|Du\|_{L^p(U)},$$

donde C es una constante que solo depende de p y U .

Esta cota nos da la posibilidad de controlar el crecimiento en $L^p(U)$ de una función $u \in W_0^{k,p}$ vía la norma de su derivada Du en $L^p(U)$.

Más aún, gracias a la definición de $\|\cdot\|_{W^{1,p}(U)}$, la cota anterior nos dice que si tenemos un subconjunto $U \subseteq \mathbb{R}$ abierto y acotado, en $W_0^{1,p}(U)$ las normas $\|\cdot\|_{W^{1,p}(U)}$ y $\|D \cdot\|_{L^p(U)}$ son equivalentes. Pasamos ahora a nuestra versión de la desigualdad.

Teorema 14. (Desigualdad de Poincaré unidimensional) Sea $N \sim \text{Normal}(0, 1)$ y sea $f \in \mathbb{D}^{1,2}$. Entonces

$$\text{Var}(f(N)) = \mathbb{E}[(Df(N))^2]$$

Demostración. Supongamos que $f \in \mathcal{S}$ y para $f \in \mathbb{D}^{1,2}$ utilizamos un argumento de aproximación. Dado que $f \in \mathcal{S}$, podemos intercambiar los símbolos de derivada e integral. Por lo tanto, tenemos que

$$\begin{aligned} \text{Var}(f(N)) &= \mathbb{E}[f(N)(f(N) - \mathbb{E}[f(N)])] \\ &= \mathbb{E}[f(N)(P_0 f(N) - P_\infty f(N))] \\ &= - \int_0^\infty \mathbb{E}\left[f(N) \frac{d}{dt} P_t f(N)\right] dt \\ &= \int_0^\infty \mathbb{E}[f(N) \delta D P_t f(N)] dt \\ &= \int_0^\infty \mathbb{E}[f'(N) D P_t f(N)] dt \\ &= \int_0^\infty e^{-t} \mathbb{E}[f'(N) P_t f'(N)] dt \\ &\leq e^{-t} \sqrt{\mathbb{E}[f'^2(N)]} \sqrt{\mathbb{E}[(P_t f')^2(N)]} dt \\ &\leq \mathbb{E}[f'^2(N)] \int_0^\infty e^{-t} dt \\ &= \mathbb{E}[f'^2(N)]. \end{aligned}$$

■

Nótese que esta desigualdad en términos de la norma $L^2(\gamma)$ se escribe como sigue:

$$\|f\|_{L^2(\gamma)}^2 \leq \mathbb{E}[f(N)] + \|Df\|_{L^2(\gamma)}^2, \quad \forall f \in \mathbb{D}^{1,2},$$

de manera que, si $\mathbb{E}[f(N)] = 0$, directamente recuperamos la desigualdad de Poincaré clásica, aunque en distinto contexto. Y como tal, esta desigualdad nos permite controlar el crecimiento de las funciones $f \in \mathbb{D}^{1,2}$ utilizando la norma de su derivada de Malliavin.

Concluimos esta sección con la denominada ‘relación de conmutatividad de Heisenberg’.

Teorema 15. Para $f \in \mathcal{S}$, y $p \geq 1$ se tiene que

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

En particular para $p = 1$, se tiene que

$$(D\delta - \delta D)f = f.$$

Demostración. Realizamos la prueba para el caso $p = 1$. El caso general se sigue usando inducción. Sea $f \in \mathcal{S}$ y notemos que para cualquier $x \in \mathbb{R}$,

$$\begin{aligned}
(D\delta - \delta D)f(x) &= D\delta f(x) - \delta Df(x) \\
&= D\delta f(x) - (xDf(x) - D^2f(x)) \\
&= D(\delta f(x) + Df(x)) - xDf(x) \\
&= D(xf(x)) - xDf(x) \\
&= f(x) + xDf(x) - xDf(x) \\
&= f(x),
\end{aligned}$$

en donde en la segunda y cuarta igualdad hemos usado la fórmula del lema 3. ■

2.2. Cálculo de Malliavin en el caso general.

En esta sección se presentan los operadores de Malliavin generales, que serán utilizados en lo subsecuente. Seguimos de cerca los requerimientos de [5, sección 2.1]

Denotemos por $\mathfrak{H} := L^2(\mathbb{R}_+ \times \mathbb{R})$. Ahora denotaremos por \mathcal{S} al siguiente conjunto:

$$\mathcal{S} := \{F = f(W(h_1), \dots, W(h_n)) : f \in C_b^\infty, h_1, \dots, h_n \in \mathfrak{H}\},$$

donde $W = \{W(h) : h \in \mathfrak{H}\}$ es el proceso Gaussiano isonormal sobre el cual estamos trabajando. Definimos ahora la derivada de Malliavin.

Definición 16. Sea $F \in \mathcal{S}$, y sea $p \geq 1$ un entero. La derivada de Malliavin de orden p de F es el elemento de $L^2(\Omega; \mathfrak{H}^{\odot p})$ definido por

$$D^p F := \sum_{i_1, \dots, i_p=1}^n \frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}}(W(h_1), \dots, W(h_n)) h_{i_1} \otimes \dots \otimes h_{i_p}.$$

En el caso en que $p = 1$ hablaremos simplemente de la «derivada de Malliavin».

Esta derivada cumple con una condición similar a la del caso unidimensional.

Proposición 16. Sea $q \in [1, \infty)$, y sea $p \geq 1$ un entero. Entonces el operador

$$D^p : \mathcal{S} \subseteq L^q(\Omega) \rightarrow L^q(\Omega; \mathfrak{H}^{\odot p})$$

es cerrable.

En consecuencia, para $q \in [1, \infty)$ y $p \geq 1$ un entero, definimos la norma $\|\cdot\|_{\mathbb{D}^{p,q}} : \mathcal{S} \rightarrow [0, \infty)$ dada por

$$\|F\|_{\mathbb{D}^{p,q}} := (\mathbb{E}[|F|^q] + \mathbb{E}[\|DF\|_{\mathfrak{H}}^q] + \dots + \mathbb{E}[\|D^p F\|_{\mathfrak{H}^{\otimes p}}^q])^{1/q}.$$

y denotamos por $\mathbb{D}^{p,q}$ a la cerradura del espacio \mathcal{S} con respecto a la norma anterior.

Pasamos ahora al operador de divergencia. Dado un entero $p \geq 1$, definimos el operador de divergencia de orden p como el operador adjunto del operador derivada $D^p : \mathbb{D}^{p,2} \rightarrow L^2(\Omega, \mathfrak{H})$.

Definición 17. Sea $p \geq 1$ un entero. Denotamos por $\text{Dom}(\delta^p)$ al subconjunto de $L^2(\Omega; \mathfrak{H}^{\otimes p})$ compuesto de aquellos elementos u tales que existe una constante $C > 0$ que satisface

$$|\mathbb{E} [\langle D^p F, u \rangle_{\mathfrak{H}^{\otimes p}}]| \leq c \sqrt{\mathbb{E} [F^2]}, \quad \forall F \in \mathcal{S}.$$

Argumentando de manera similar al caso unidimensional, por el teorema de representación de Riesz, existe un único elemento en $L^2(\Omega)$, denotado por $\delta^p(u)$ tal que cumple la relación

$$\mathbb{E} [\langle D^p F, u \rangle_{\mathfrak{H}^{\otimes p}}] = \mathbb{E} [F \delta^p(u)], \quad \forall f \in \mathcal{S}.$$

Llegamos así a la siguiente definición.

Definición 18. Si $u \in \text{Dom}(\delta^p)$, entonces $\delta^p(u)$ es el único elemento de $L^2(\Omega)$ caracterizado por la siguiente fórmula:

$$\mathbb{E} [\langle D^p F, u \rangle_{\mathfrak{H}^{\otimes p}}] = \mathbb{E} [F \delta^p(u)], \quad \forall f \in \mathcal{S}.$$

Este operador se denomina el operador de divergencia múltiple de orden p . Cuando $p = 1$ simplemente lo llamamos el «operador de divergencia», escribiendo δ en lugar de δ^1 . La fórmula anterior se suele llamar «fórmula de integración por partes».

El siguiente resultado nos permite «sacar» una variable aleatoria escalar dentro del operador de divergencia δ . Este resultado se puede consultar en [7, proposición 2.5.4], sin embargo, enunciamos una versión general que se encuentra en [1, lema 1].

Teorema 16. Sean $p, p' > 1$ tales que $\frac{1}{p} + \frac{1}{p'} = 1$. Supongamos que $F \in \mathbb{D}^{1,p'}$, y sea $u \in \text{Dom}(\delta)$ tal que $u \in L^p(\Omega; \mathfrak{H})$ y $\delta(u) \in L^p(\Omega)$. Entonces $Fu \in \text{Dom}(\delta)$ y se cumple la igualdad

$$\delta(Fu) = F\delta(u) - \langle DF, u \rangle_{\mathfrak{H}}. \quad (2.1)$$

Tenemos también la propiedad de conmutatividad de Heisenberg en este caso. Dicha propiedad es complemente análoga al caso unidimensional.

Teorema 17. Se cumple que

$$D\delta(u) - \delta(Du) = u.$$

Finalmente, es de destacar una propiedad muy interesante del operador de divergencia: cuando estamos en el caso en que $\mathfrak{H} = L^2(\mathbb{R}_+ \times \mathbb{R})$, el operador delta coincide con la integral de Itô-Walsh.

Teorema 18. Sea u un proceso adaptado y cuadrado-integrable. Entonces $u \in \text{Dom}(\delta)$ y se cumple que

$$\delta(v) = \int_{\mathbb{R}_+ \times \mathbb{R}} v(s, y) W(ds, dy).$$

Capítulo 3

Aplicaciones a Teoremas Límite y estudio de densidades

En este capítulo se visita la aplicación de los conceptos del cálculo de Malliavin al estudio de Teoremas límite y al estudio de densidades. Una herramienta clave en esta tarea es el conocido como Método de Stein. Dicho método, desarrollado por Charles Stein en la década de 1970, es una herramienta bastante popular para evaluar la distancia entre las leyes de dos leyes de probabilidad. Lo anterior es posible lograrlo estudiando operadores diferenciales involucrados con las leyes de probabilidad en cuestión.

3.1. Método de Stein

Dentro del mundo de la probabilidad, la distribución normal juega un papel importante debido a la gran cantidad de propiedades que la misma presenta. Entre ellas, dicha distribución está determinada por sus momentos. Esto hace posible caracterizar la distribución en términos de los mismos. Ésta propiedad, la cual está codificada en el conocido como ‘lema de Stein’, forma parte importante de este texto.

Comenzamos dando una prueba de que la ley de una variable Gaussiana estándar está caracterizada por sus momentos. Recordemos que, dada una medida ν en \mathbb{R} , el n -ésimo momento de la misma, cuando existe, está dado por

$$m_n(\nu) := \int_{\mathbb{R}} x^n d\nu(x), \quad n \geq 0.$$

Lema 4. Sea γ la distribución gaussiana estándar. Dicha distribución está caracterizada por sus momentos.

Demostración. Debemos mostrar que, si existe otra medida definida en \mathbb{R} con todos sus momentos definidos, digamos μ , y tales momentos coinciden con los de γ , entonces $\mu = \gamma$.

Sea entonces μ una medida como antes. Es suficiente mostrar que la función característica de ambas medidas coincide en todo \mathbb{R} , esto es, que

$$\int_{\mathbb{R}} e^{itx} d\mu(x) = \int_{\mathbb{R}} e^{itx} d\gamma(x), \quad \forall t \in \mathbb{R}.$$

Utilizando la forma de Lagrange del residuo en el teorema de Taylor y la desigualdad de Cauchy-Schwarz, tenemos para cualquier $n \geq 1$ que

$$\begin{aligned} \left| \int_{\mathbb{R}} e^{itx} d\nu(x) - \int_{\mathbb{R}} e^{itx} d\gamma(x) \right| &\leq \int_{\mathbb{R}} \left| e^{itx} - \sum_{k=0}^n \frac{(itx)^k}{k!} \right| d\nu(x) + \int_{\mathbb{R}} \left| e^{itx} - \sum_{k=0}^n \frac{(itx)^k}{k!} \right| d\gamma(x) \\ &= \int_{\mathbb{R}} \left| \frac{(it)^{n+1} e^{it\xi_L} x^{n+1}}{(n+1)!} \right| d\nu(x) + \int_{\mathbb{R}} \left| \frac{(it)^{n+1} e^{it\eta_L} x^{n+1}}{(n+1)!} \right| d\gamma(x) \\ &\leq \int_{\mathbb{R}} \frac{|tx|^{n+1}}{(n+1)!} d\nu(x) + \int_{\mathbb{R}} \frac{|tx|^{n+1}}{(n+1)!} d\gamma(x) \\ &\leq \left(\int_{\mathbb{R}} \frac{|tx|^{2n+2}}{(n+1)!^2} d\nu(x) \right)^{1/2} + \left(\int_{\mathbb{R}} \frac{|tx|^{2n+2}}{(n+1)!^2} d\gamma(x) \right)^{1/2} \\ &= \left(\frac{|t|^{2n+2}}{(n+1)!^2} m_{2n+2}(\nu) \right)^{1/2} + \left(\frac{|t|^{2n+2}}{(n+1)!^2} m_{2n+2}(\gamma) \right)^{1/2} \\ &= 2 \left(\frac{|t|^{2n+2}}{(n+1)!^2} m_{2n+2}(\gamma) \right)^{1/2}. \end{aligned}$$

Recordando que para la distribución Gaussiana se cumple que $m_{2n}(\gamma) = \frac{(2n)!}{2^n n!}$, y usando cotas superior e inferior de la fórmula de Stirling, tenemos que

$$\begin{aligned} 2 \left(\frac{|t|^{2n+2}}{(n+1)!^2} m_{2n+2}(\gamma) \right)^{1/2} &= 2|t|^{n+1} \left(\frac{(2n+2)!}{2^{n+1}(n+1)!^3} \right)^{1/2} \\ &\leq 2|t|^{n+1} \left(\frac{\sqrt{4\pi(n+1)} \left(\frac{2(n+1)}{e} \right)^{2n+2} e^{-12 \cdot 2(n+1)}}{2^{n+1} \left(\sqrt{2\pi(n+1)} \left(\frac{n+1}{e} \right)^{n+1} e^{-12(n+1)+1} \right)^3} \right)^{1/2} \\ &\leq 2|t|^{n+1} \left(\frac{2^n e^{13(n+1)+3}}{\pi(n+1)^{n+2}} \right)^{1/2} \\ &= \frac{K_t^{\frac{13}{2}(n+1)}}{(n+1)^{n/2+1}}, \end{aligned}$$

donde K_t es una constante dependiente sólo de t . Pero esta última expresión tiende a cero cuando n tiende a infinito. Concluimos que las funciones características de ambas medidas coinciden y por lo tanto, son equivalentes. ■

Ahora enunciamos y demostramos el conocido como ‘lema de Stein’.

Teorema 19. (Lema de Stein) Sea X una variable aleatoria. X tiene una distribución normal estándar si y sólo si, para cualquier función f derivable y tal que $f' \in L^1(\gamma)$, se tiene que

1. $\mathbb{E}[Xf(X)] < \infty$ y $\mathbb{E}[f'(X)] < \infty$.
2. $\mathbb{E}[Xf(X)] = \mathbb{E}[f'(X)]$.

Demostración. Si directamente $X \sim \text{Normal}(0, 1)$, cualquier función que cumpla las condiciones del enunciado cumple las del lema 1, el cual precisamente estipula la relación anterior.

Supongamos ahora que X satisface la relación $\mathbb{E}[Xf(X)] = \mathbb{E}[f'(X)]$ para cualquier función f con las propiedades del enunciado. Si definimos $f_n(x) = x^n$ para cualquier $n \geq 1$, tenemos que la variable X tiene todos sus momentos definidos, y además

$$\mathbb{E}[X^{n+1}] = \mathbb{E}[Xf(X)] = \mathbb{E}[f'(X)] = n\mathbb{E}[X^{n-1}],$$

de tal forma que los momentos de la variable X coinciden con los de una variable normal estándar. Dado que la distribución normal está caracterizada por sus momentos según vimos en el resultado previo, se sigue que $X \sim \text{Normal}(0, 1)$. ■

El resultado anterior es interesante por varios motivos. Uno de ellos es que nos permite caracterizar plenamente la distribución normal a partir de funciones y sus derivadas aplicadas a una variable aleatoria que siga dicha distribución. Por otro lado, un instante de reflexión nos puede llevar a la siguiente pregunta.

Supongamos que ahora X es una variable aleatoria y que para una cierta clase suficientemente rica de funciones suaves, se cumple que

$$\mathbb{E}[Xf(X) - f'(X)] \approx 0.$$

En virtud del lema de Stein, ¿es posible decir que la distribución de X está cerca (en algún sentido) de la distribución normal estándar?

La pregunta anterior se conoce como la *heurística de Stein*, y la respuesta es afirmativa en algunos casos. Para llegar a una formulación explícita de dicha respuesta, debemos introducir la idea de ‘distancia entre dos medidas de probabilidad’ y también la famosa ‘ecuación de Stein’ asociada a alguna función $h \in L^1(\gamma)$.

La definición de distancia se otorga para variables que tomen valores en \mathbb{R}^d . Comenzamos definiendo la noción de distancia entre dos variables aleatorias. Recordamos la noción de clase separante de funciones.

Definición 19. En un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, sea \mathcal{H} un conjunto de funciones borelianas complejas con valores en \mathbb{R}^d , para algún $d \geq 1$. Decimos que \mathcal{H} es una familia de funciones separantes si para cualesquiera dos variables aleatorias F y G con valores en \mathbb{R}^d

tales que $h(F)$ y $h(G)$ son elementos de $L^1(\Omega)$ y $\mathbb{E}[h(F)] = \mathbb{E}[h(G)]$, para cualquier elemento $h \in \mathcal{H}$ implica que F y G tienen la misma ley.

Ejemplos clásicos de familias \mathcal{H} separantes son el conjunto de funciones borelianas continuas y acotadas, el conjunto de funciones indicadoras de los borelianos de \mathbb{R}^d , o el conjunto de funciones exponenciales de la forma $\mathcal{H} = \{e^{i\langle v, \cdot \rangle} : v \in \mathbb{R}^d\}$, donde el producto interno anterior es el producto interno usual de \mathbb{R}^d .

Definimos a continuación, y de manera rigurosa, la noción de distancia entre las leyes de dos variables aleatorias.

Definición 20. Sea \mathcal{H} una clase separante de funciones, y sean F y G dos variables aleatorias con valores en \mathbb{R}^d y tales que $h(F)$ y $h(G)$ están en $L^1(\Omega)$ para cualquier $h \in \mathcal{H}$. Definimos la distancia inducida por \mathcal{H} entre las distribuciones de F y G como

$$d_{\mathcal{H}}(F, G) := \sup_{h \in \mathcal{H}} \{|\mathbb{E}[h(F)] - \mathbb{E}[h(G)]|\}.$$

Es sencillo demostrar que para una clase de funciones separantes \mathcal{H} , el mapeo $d_{\mathcal{H}}$ definido anteriormente en efecto define una métrica en cierto subconjunto de las medidas de probabilidad sobre \mathbb{R}^d . Es pertinente hacer hincapié en que la distancia inducida entre las leyes de variables depende de la clase \mathcal{H} . Ello resulta en una amplia variedad de opciones con las cuales considerar la distancia entre dos distribuciones de probabilidad. Algunas de las más usadas se encuentran a continuación.

- **La distancia de Kolmogorov.** Esta distancia entre dos variables F y G , denotada por $d_{\text{Kol}}(F, G)$, se obtiene utilizando a la familia

$$\mathcal{H} = \{h : \mathbb{R}^d \rightarrow \mathbb{R} : h(x_1, \dots, x_d) = \mathbb{1}_{(-\infty, z_1]}(x_d) \cdot \dots \cdot \mathbb{1}_{(-\infty, z_d]}(x_d), \quad z_1, \dots, z_d \in \mathbb{R}\},$$

de manera que dicha distancia está dada por la expresión

$$d_{\text{Kol}}(F, G) = \sup_{z_1, \dots, z_d \in \mathbb{R}} |\mathbb{P}(F \in (-\infty, z_1] \times \dots \times (-\infty, z_d]) - \mathbb{P}(G \in (-\infty, z_1] \times \dots \times (-\infty, z_d])|.$$

- **La distancia de variación total.** Dicha distancia entre las leyes de las variables F y G se define utilizando la clase

$$\mathcal{H} = \{h : \mathbb{R}^d \rightarrow \mathbb{R} : h(x_1, \dots, x_d) = \mathbb{1}_B(x_1, \dots, x_d), \quad B \in \mathcal{B}(\mathbb{R}^d)\},$$

y dicha distancia se denota por $d_{\text{TV}}(F, G)$. Su expresión explícita es

$$d_{\text{TV}}(F, G) = \sup_{B \in \mathcal{B}(\mathbb{R}^d)} |\mathbb{P}(F \in B) - \mathbb{P}(G \in B)|.$$

Observamos directamente que $d_{\text{Kol}}(F, G) \leq d_{\text{TV}}(F, G)$.

▪ **La distancia de Wasserstein.** Usando la clase de funciones separantes

$$\mathcal{H} = \left\{ h : \mathbb{R}^d \rightarrow \mathbb{R} : \|h\|_{\text{Lip}} \leq 1 \right\},$$

donde claramente $\|\cdot\|_{\text{Lip}}$ denota la norma Lipschitz en \mathbb{R}^d . Esta distancia se denota como $d_W(F, G)$.

Armados con la noción de distancia entre dos leyes de probabilidad, pasamos ahora a estudiar brevemente las ecuaciones de Stein. Dichas ecuaciones nos permiten codificar la distancia entre la ley de una variable F y una variable aleatoria normal N , en virtud del lema de Stein 19. De manera rigurosa, tenemos la siguiente definición.

Definición 21. Sea N una variable gaussiana estándar y sea $h : \mathbb{R} \rightarrow \mathbb{R}$ una función boreliana tal que $\mathbb{E}[h(N)] < \infty$. La ecuación de Stein asociada con h es la ecuación diferencial ordinaria

$$f'(x) - xf(x) = h(x) - \mathbb{E}[h(N)]. \quad (3.1)$$

Diremos que f es una solución a la ecuación (3.1) si dicha función es absolutamente continua y tal que existe función equivalente a la derivada f' que satisface (3.1) para cualquier $x \in \mathbb{R}$.

Es de notar que dicha ecuación diferencial es sencilla de resolver. Este es el contenido del siguiente resultado.

Teorema 20. Las soluciones a la ecuación (3.1) son de la forma

$$f(x) = ce^{x^2/2} + e^{x^2/2} \int_{-\infty}^x (h(y) - \mathbb{E}[h(N)])e^{-y^2/2} dy, \quad \forall x \in \mathbb{R},$$

donde $c \in \mathbb{R}$ es constante. En particular, si escribimos

$$f_h(x) := e^{x^2/2} \int_{-\infty}^x (h(y) - \mathbb{E}[h(N)])e^{-y^2/2} dy,$$

entonces f_h es la única solución a la ecuación de Stein (3.1) que satisface la condición

$$\lim_{x \rightarrow \infty} e^{-x^2/2} f(x) = 0.$$

Demostración. Supongamos que h cumple las hipótesis del enunciado. Partiendo de las igualdades

$$\frac{d}{dx} \left(e^{-x^2/2} f(x) \right) = -xe^{-x^2/2} f(x) + e^{-x^2/2} f'(x) = e^{-x^2/2} (f'(x) - xf(x)),$$

la ecuación de Stein (3.1) se puede reformular como

$$e^{x^2/2} \frac{d}{dx} \left(e^{-x^2/2} f(x) \right) = h(x) - \mathbb{E}[h(N)].$$

Usando el método de factor integrante, es claro que las soluciones a la ecuación (3.1) tienen la forma

$$f(x) = ce^{x^2/2} + x^{2/2} \int_{-\infty}^x (h(y) - \mathbb{E}[h(N)])e^{-y^2/2}dy, \quad \forall x \in \mathbb{R}.$$

Para el comportamiento en $\pm\infty$ de las solución f_h , usando la función $g(x) = |h(y) - \mathbb{E}[h(N)]|e^{-y^2/2}$, la cual es integrable e independiente de $x \in \mathbb{R}$, por el teorema de convergencia dominada, tenemos para el límite en $+\infty$ que

$$\begin{aligned} \lim_{x \rightarrow \infty} \int_{-\infty}^x (h(y) - \mathbb{E}[h(N)])e^{-y^2/2}dy &= \int_{\mathbb{R}} \lim_{x \rightarrow \infty} (h(y) - \mathbb{E}[h(N)])e^{-y^2/2} \mathbf{1}_{(-\infty, x]}(y)dy \\ &= \int_{\mathbb{R}} h(y)e^{-y^2/2}dy - \mathbb{E}[h(N)] \int_{\mathbb{R}} e^{-y^2/2}dy \\ &= \sqrt{2\pi}\mathbb{E}[h(N)] - \sqrt{2\pi}\mathbb{E}[h(N)] \\ &= 0, \end{aligned}$$

mientras que para el límite en $-\infty$ tenemos que

$$\lim_{x \rightarrow -\infty} \int_{-\infty}^x (h(y) - \mathbb{E}[h(N)])e^{-y^2/2}dy = \int_{\mathbb{R}} \lim_{x \rightarrow -\infty} (h(y) - \mathbb{E}[h(N)])e^{-y^2/2} \mathbf{1}_{(-\infty, x]}(y)dy = 0.$$

Se sigue entonces que si f es una solución como las anteriormente mencionadas, entonces $e^{-x^2/2}f(x) \xrightarrow{x \rightarrow \pm\infty} 0$ si y solo si $c = 0$. ■

Contando ahora con estas herramientas, podemos evaluar de manera rigurosa la distancia que hay entre la ley de una variable X y una variable normal estándar N . Para ello, consideremos una función $h : \mathbb{R} \rightarrow \mathbb{R}$ que cumple que $\mathbb{E}[|h(X)|] < \infty$ y $\mathbb{E}[|h(N)|] < \infty$. Sea f_h la solución única del teorema 20. Tomando esperanza a ambos lados de la ecuación de Stein, tenemos que f_h cumple que

$$\mathbb{E}[h(X)] - \mathbb{E}[h(N)] = \mathbb{E}[f'_h(X) - Xf_h(X)]. \quad (3.2)$$

En particular, notamos que para una clase de funciones adecuada \mathcal{H} ,

$$d_{\mathcal{H}}(X, N) = \sup_{h \in \mathcal{H}} |\mathbb{E}[h(X)] - \mathbb{E}[h(N)]| = \sup_{h \in \mathcal{H}} |\mathbb{E}[f'_h(X) - Xf_h(X)]|.$$

Es de destacar que en la expresión anterior, la distancia entre las leyes de las variables X y N están codificadas en la cantidad de la derecha, la cual no involucra la ley Gaussiana estándar de manera directa. Por otro lado, es deseable desarrollar algunas cotas para la norma de las soluciones f_h , dependiendo del tipo de funciones $h \in \mathcal{H}$ que estemos considerando. Dependiendo de dicha clase, podremos obtener distintas cotas para las distancias en las que estemos interesados. A dichas cotas se les conoce como *cotas de Stein*.

En nuestro caso, nos vamos a enfocar en dos familias de funciones distintas. La primera de ellas es la familia \mathcal{H} que origina la distancia de Kolmogorov, mientras que en el segundo caso consideramos una familia específica que nos será de utilidad en el capítulo 4.

En cuanto a la distancia de Kolmogorov, tenemos el siguiente teorema que nos da una cota universal explícita de las normas de las soluciones a la ecuación de Stein asociada a la familia $\mathcal{H} = \{h : \mathbb{R} \rightarrow \mathbb{R} : h(x) = \mathbb{1}_{(-\infty, z]}(x), \quad z \in \mathbb{R}\}$.

Teorema 21. Sea $z \in \mathbb{R}$ y consideremos a la función $h = \mathbb{1}_{(-\infty, z]} \in \mathcal{H}$ asociada. Denotemos por f_z a la única solución de la ecuación de Stein asociada a la función h anterior. Se tiene la siguiente cota.

$$\|f_z\|_\infty \leq \frac{\sqrt{2\pi}}{4} \quad \text{y} \quad \|f'_z\|_\infty \leq 1.$$

En particular, para N una variable aleatoria normal estándar, y para cualquier variable integrable X , se tiene que

$$d_{\text{Kol}}(X, N) \leq \sup_{f \in \mathcal{F}_{\text{Kol}}} |\mathbb{E}[f'(X)] - \mathbb{E}[Xf(X)]|,$$

en donde

$$\mathcal{F}_{\text{Kol}} := \left\{ f : \mathbb{R} \rightarrow \mathbb{R} \text{ abs. cont. : } f \text{ derivable salvo un conjunto finito, } \|f\|_\infty \leq \frac{\sqrt{2\pi}}{4}, \|f'\|_\infty \leq 1 \right\},$$

y donde el supremo anterior se entiende como la cantidad

$$\sup |\mathbb{E}[g(X)] - \mathbb{E}[Xf(X)]|,$$

donde el supremo se toma sobre todos los pares de funciones borelianas (f, g) , con f derivable salvo un conjunto finito de puntos, y g es una versión de f' tal que $\|g\|_\infty \leq 1$.

Una guía para la demostración de este resultado se puede encontrar en la sección 3.4 de [7]. Concluimos esta sección con el siguiente resultado, el cual nos da cotas explícitas para una familia de funciones específicas.

Teorema 22. Consideremos la familia de funciones $\mathcal{H} = \{h : \mathbb{R} \rightarrow \mathbb{R} : h(z) = \mathbb{1}_{\{z > x\}}z, \quad x \in \mathbb{R}\}$. Entonces la solución f_h a la ecuación de Stein (3.1) satisface

$$|f'_h(x)| \leq C(|x| + 1), \quad \forall x \in \mathbb{R},$$

en donde C es una constante.

Nuevamente este resultado se enuncia sin demostración. Una versión general del mismo, así como su demostración, se encuentran en el apéndice A de [2].

Los dos resultados anteriores se encuentran en el núcleo de las técnicas que se conocen en su conjunto como *método de Stein*, y que son ampliamente usadas para deducir resultados similares al teorema central del límite, además de deducir tasas de convergencia para los mismos. Ejemplo de lo anterior es el célebre teorema de Berry-Essen, el cual cuantifica la velocidad de convergencia de una sucesión i.i.d de variables con media cero, varianza 1 y

tercer momento absoluto finito, a la distribución normal estándar. Es decir, es una versión cuantificable del clásico teorema del límite central.

Es posible otorgar una demostración del mismo utilizando el método de Stein, y más específicamente el teorema 21. Para el enunciado y demostración del mismo, nos referimos a la sección 3.7 de [7].

3.2. Existencia de densidades y estimación de la distancia uniforme

En preparación para el capítulo 4, en esta sección estudiamos brevemente un par de resultados. El primero de ellos nos garantiza la existencia y unicidad de una variable aleatoria unidimensional bajo ciertas hipótesis, además de proveer una fórmula para la misma. También enunciamos un resultado que consiste en una cota para la distancia uniforme entre la densidad de una variable aleatoria y la densidad de una variable normal estándar, siendo este último un resultado medular en el capítulo 4. Es de destacar que estos resultados son consecuencias del teorema 16 y ambos, por orden de aparición, se pueden encontrar en [1, proposición 1] y en [5, teorema 3.2].

Comenzamos directamente con el primer resultado. La demostración del mismo se omite en este texto.

Teorema 23. Sean F una variable aleatoria en el espacio $\mathbb{D}^{1,1}$ y \mathfrak{H} un espacio de Hilbert. Supongamos que u es una variable en $L^1(\Omega; \mathfrak{H})$ tal que $D_u F \neq 0$ casi seguramente. Supongamos también que $\frac{u}{D_u F} \in \text{Dom}(\delta)$. Entonces la ley de F tiene una densidad continua y acotada dada por

$$f_F(x) = \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \delta \left(\frac{u}{D_u F} \right) \right]. \quad (3.3)$$

Es de nuestro interés saber en qué caso es cierto que $\frac{u}{D_u F} \in \text{Dom}(\delta)$. El siguiente lema nos da una respuesta.

Lema 5. Sean $p, p' > 1$ tales que $\frac{1}{p} + \frac{1}{p'} = 1$, y sea $F \in \mathbb{D}^{2,1}$. Supongamos que $u \in \text{Dom}(\delta)$. Si se cumplen las siguientes condiciones

1. $u \in L^p(\Omega; \mathfrak{H})$,
2. $\delta(u) \in L^p(\Omega)$,
3. $(D_u F)^{-1} \in L^{p'}(\Omega)$,
4. $(D_u F)^{-2} (\|D^2 F\|_{\mathfrak{H} \otimes 2} \|u\|_{\mathfrak{H}} + \|Du\|_{\mathfrak{H}} \|DF\|_{\mathfrak{H}}) \in L^{p'}(\Omega)$,

entonces $\frac{u}{D_u F} \in \text{Dom}(\delta)$ y en particular, se cumple la ecuación (2.1).

Para una prueba de este lema, se puede consultar [1, lema 3, lema 4].

Estamos interesados en aplicar tanto el lema anterior como el teorema 23 en condiciones que aparecerán en el capítulo 4. Por lo tanto, y siguiendo de cerca a [5, sección 3], tenemos el siguiente resultado.

Teorema 24. Sean $u \in \mathbb{D}^{1,6}(\Omega; \mathfrak{H})$ y definamos la variable $F := \delta(u)$. Supongamos que $F \in \mathbb{D}^{2,6}$, con $\mathbb{E}[F] = 0$, $\mathbb{E}[F^2] = 1$ y $(D_u F)^{-1} \in L^4(\Omega)$. Entonces $u/D_u F \in \text{Dom}(\delta)$, la variable F admite una densidad continua y acotada $f_F(x)$ y la misma cumple la siguiente desigualdad

$$\sup_{x \in \mathbb{R}} |f_F(x) - \Phi(x)| \leq (\|F\|_4 \| (D_v F)^{-1} \|_4 + 2) \|1 - D_v F\|_2 + \| (D_v F)^{-1} \|^2 \|D_v (D_v F)\|_2, \quad (3.4)$$

donde $\Phi(x)$ es la función de densidad de una variable normal estándar.

Demostración. Demostramos primero que $\frac{u}{D_u F} \in \text{Dom}(\delta)$. Buscamos satisfacer las hipótesis del lema 5. Para ello usamos $p = p' = 2$. Claramente $p, p' > 1$ y se cumple que $\frac{1}{p} + \frac{1}{p'} = 1$. Por otro lado, por hipótesis $u \in \text{Dom}(\delta)$ y además, dado que en nuestro caso la variable F es por definición la divergencia de u , de la hipótesis y de la contención entre los espacios $\mathbb{D}^{k,q}$ se sigue que $F \in \mathbb{D}^{2,6} \subseteq \mathbb{D}^{2,1}$. Resta mostrar que las cuatro condiciones del lema 5 se satisfacen.

1. Dado que $u \in \mathbb{D}^{1,6}(\Omega; \mathfrak{H}) \subseteq \mathbb{D}^{1,2}(\Omega; \mathfrak{H})$, y que por construcción este último espacio es un subconjunto de $L^2(\Omega; \mathfrak{H})$, se cumple la primera condición.
2. Por hipótesis $\delta(u) = F \in \mathbb{D}^{2,6}$, y este es subconjunto de $L^2(\Omega)$, por lo que la segunda condición también se cumple.
3. El enunciado nos indica que $(D_u F)^{-1} \in L^4(\Omega)$. Como la medida del espacio es finita, se da la contención $L^4(\Omega) \subseteq L^2(\Omega)$ y la tercera condición se satisface.
4. Denotemos $a := (\|D^2 F\|_{\mathfrak{H} \otimes 2} \|u\|_{\mathfrak{H}} + \|Du\|_{\mathfrak{H}} \|DF\|_{\mathfrak{H}})$. Obsérvese que dichas cantidades están bien definidas gracias a las hipótesis. Notamos que

$$\|a(D_u F)^{-2}\|_2^2 = \int_{\Omega} (a(D_u F)^{-2})^2 d\mathbb{P} = a^2 \int_{\Omega} ((D_u F)^{-1})^4 d\mathbb{P} = a^2 \|(D_u F)^{-1}\|_4^4 < \infty.$$

Concluimos que la última condición se satisface.

Por lo tanto, $u/D_u F \in \text{Dom}(\delta)$. Ahora bien, el resto de hipótesis del teorema 23 se satisfacen, pues se tiene que $F \in \mathbb{D}^{2,6} \subseteq \mathbb{D}^{1,1}$. Además, $u \in \mathbb{D}^{1,6} \subseteq \mathbb{D}^{1,1}$ y este último es subconjunto de $L^1(\Omega; \mathfrak{H})$. Finalmente si fuera el caso que $D_u F = 0$ en un conjunto medible de probabilidad positiva $A \subseteq \Omega$, tendríamos que

$$\|(D_u F)^{-1}\|_1 = \int_{\Omega} |(D_u F)^{-1}| d\mathbb{P} \geq \int_A |(D_u F)^{-1}| d\mathbb{P} = +\infty,$$

lo cual es una contradicción ya que $(D_u F)^{-1} \in L^4(\Omega)$, que es subconjunto de $L^1(\Omega)$.

Satisfechas las hipótesis del teorema 23, tenemos garantizada la existencia de una densidad continua y acotada para la variable $F = \delta(v)$, la cual está dada por la fórmula (3.3). Ahora buscamos una cota uniforme sobre $x \in \mathbb{R}$ para las siguientes cantidades

$$|f_F(x) - \phi(x)| = \left| \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \delta \left(\frac{u}{D_u F} \right) \right] - \mathbb{E} [\mathbb{1}_{\{N > x\}} N] \right|, \quad (3.5)$$

donde $N \sim \text{Normal}(0, 1)$. En virtud de que se satisfacen las hipótesis del lema 23, en la fórmula (3.3) podemos usar la ecuación (2.1) y obtenemos que, para $x \in \mathbb{R}$ arbitrario,

$$\begin{aligned} f_F(x) &= \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \delta \left(\frac{u}{D_u F} \right) \right] = \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \left(\frac{1}{D_u F} \delta(u) - \left\langle D \left(\frac{1}{D_u F} \right), u \right\rangle \right) \right] \\ &= \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \frac{F}{D_u F} \right] + \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \left(\frac{1}{(D_u F)^2} \right) D_u D_u F \right], \end{aligned}$$

donde hemos hecho uso de que $F = \delta(u)$ y de la regla de la cadena en la segunda igualdad. Incrustando esta última igualdad en la fórmula (3.5), y subsecuentemente sumando y restando el término $\mathbb{E} [\mathbb{1}_{\{F > x\}} F]$,

$$\begin{aligned} |f_F(x) - \Phi(x)| &= \left| \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \frac{F}{D_u F} \right] + \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \frac{D_u D_u F}{(D_u F)^2} \right] - \mathbb{E} [\mathbb{1}_{\{N > x\}} N] \right| \\ &= \left| \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \frac{F}{D_u F} \right] - \mathbb{E} [\mathbb{1}_{\{F > x\}} F] + \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \frac{D_u D_u F}{(D_u F)^2} \right] \right. \\ &\quad \left. + \mathbb{E} [\mathbb{1}_{\{F > x\}} F] - \mathbb{E} [\mathbb{1}_{\{N > x\}} N] \right| \\ &\leq \left| \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \left(\frac{F}{D_u F} - F \frac{D_u F}{D_u F} \right) \right] \right| + \left| \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \frac{D_u D_u F}{(D_u F)^2} \right] \right| \\ &\quad + \left| \mathbb{E} [\mathbb{1}_{\{F > x\}} F] - \mathbb{E} [\mathbb{1}_{\{N > x\}} N] \right| \\ &\leq \mathbb{E} \left[\left| \frac{F(1 - D_u F)}{D_u F} \right| \right] + \mathbb{E} \left[\left| \frac{D_u D_u F}{(D_u F)^2} \right| \right] + |\mathbb{E} [\mathbb{1}_{\{F > x\}} F] - \mathbb{E} [\mathbb{1}_{\{N > x\}} N]|. \quad (3.6) \end{aligned}$$

Ahora, en la desigualdad (3.6) usamos sucesivamente la desigualdad de Hölder en el primer y segundo términos, de manera que

$$\mathbb{E} \left[\left| \frac{F(1 - D_u F)}{D_u F} \right| \right] \leq \|F(D_u F)\|_2 \|1 - D_u F\|_2 \leq \|F\|_4 \|(D_u F)^{-1}\|_4 \|1 - D_u F\|_2, \quad (3.7)$$

y también

$$\mathbb{E} \left[\left| \frac{D_u D_u F}{(D_u F)^2} \right| \right] \leq \|D_u D_u F\|_2 \|(D_u F)^{-2}\|_2 = \|D_u D_u F\|_2 \|(D_u F)^{-1}\|_4^2. \quad (3.8)$$

Finalmente, estudiamos el tercer término de la desigualdad (3.6). Aquí es donde utilizamos el «método de Stein». Observemos que, para $x \in \mathbb{R}$,

$$|\mathbb{E} [\mathbf{1}_{\{F > x\}} F] - \mathbb{E} [\mathbf{1}_{\{N > x\}} N]| = |\mathbb{E} [h(F)] - \mathbb{E} [h(N)]|,$$

donde $h(z) = \mathbf{1}_{\{z > x\}} z$ es una función perteneciente a la familia \mathcal{H} correspondiente al teorema 22. En virtud de la ecuación (3.2), la igualdad anterior se convierte en

$$|\mathbb{E} [\mathbf{1}_{\{F > x\}} F] - \mathbb{E} [\mathbf{1}_{\{N > x\}} N]| = |\mathbb{E} [f'_h(F)] - \mathbb{E} [F f_h(F)]|. \quad (3.9)$$

En este punto usamos nuevamente que $F = \delta(u)$, y como tal, cumple la fórmula de integración por partes. Por lo tanto, tomando como variable a $f_h(F)$ y usando regla de la cadena se tienen las siguientes igualdades

$$\mathbb{E} [F f_h(F)] = \mathbb{E} [\delta(u) f_h(F)] = \mathbb{E} [\langle D(f_h(F)), u \rangle] = \mathbb{E} [f'_h(F) \langle DF, u \rangle] = \mathbb{E} [f'_h(F) D_u F].$$

Insertando la igualdad anterior en (3.9) se sigue que

$$|\mathbb{E} [\mathbf{1}_{\{F > x\}} F] - \mathbb{E} [\mathbf{1}_{\{N > x\}} N]| = |\mathbb{E} [f'_h(F)] - \mathbb{E} [f'_h(F) D_u F]| = |\mathbb{E} [f'_h(F)(1 - D_v F)]|.$$

Usamos nuevamente la desigualdad de Hölder en el último termino de la ecuación anterior, obtenemos

$$|\mathbb{E} [f'_h(F)(1 - D_v F)]| \leq \|f'_h(F)\|_2 \|1 - D_v F\|_2. \quad (3.10)$$

Finalmente, usamos el teorema 22, en 3.10, por lo que

$$|\mathbb{E} [\mathbf{1}_{\{F > x\}} F] - \mathbb{E} [\mathbf{1}_{\{N > x\}} N]| \leq C(|F| + 1) \|1 - D_v F\|_2 \quad (3.11)$$

Finalmente, tenemos que $|F| \leq 1$ y $C \leq 1$, por lo que 3.11 se convierte en

$$|\mathbb{E} [\mathbf{1}_{\{F > x\}} F] - \mathbb{E} [\mathbf{1}_{\{N > x\}} N]| \leq 2 \|1 - D_v F\|. \quad (3.12)$$

Combinando las desigualdades 3.7, 3.8 y 3.11, tenemos para $x \in \mathbb{R}$,

$$\begin{aligned} |f_F(x) - \Phi(x)| &\leq \|F\|_4 \|(D_v F)^{-1}\|_4 \|1 - D_u F\|_2 + \|D_u D_u F\|_2 \|(D_u F)^{-1}\|_4^2 + 2 \|1 - D_v F\| \\ &= (\|F\|_4 \|(D_v F)^{-1}\|_4 + 2) \|1 - D_v F\|_2 + \|(D_v F)^{-1}\|^2 \|D_v (D_v F)\|_2. \end{aligned}$$

Tomando supremo obtenemos la desigualdad (3.4). ■

Tal y como se comentó al inicio de la sección, el resultado anterior resulta ser la cota principal desde la cual se deducirá la tasa de convergencia de las densidades de los promedios espaciales de la solución a la SHE, a la densidad normal.

Hasta aquí llega nuestro estudio del método de Stein y de la existencia de las densidades.

Capítulo 4

Propiedades espaciales de la Ecuación estocástica del calor

En este capítulo presentaremos la parte central de la aplicación de las técnicas de Malliavin-Stein y SPDE's al estudio de la convergencia de las densidades de los promedios espaciales de la solución de la ecuación estocástica del calor, a la densidad de una variable aleatoria gaussiana estándar.

Consideremos la siguiente ecuación diferencial parcial estocástica:

$$\partial u_t = \frac{1}{2} \partial_{xx} u + \sigma(u) \dot{W}, \quad \forall x \in \mathbb{R}, \forall t > 0. \quad (4.1)$$

cuya condición inicial es $u(0, x) := u_0(x)$, y \dot{W} es un ruido blanco en $[0, \infty) \times \mathbb{R}$, esto es, tanto en el espacio como en el tiempo.

Del capítulo 1, tenemos garantizada la existencia y unicidad de una solución mild $u(t, x)$ a la ecuación 1, suponiendo que σ es una función Lipschitz, y que u_0 es una medida signada que satisface la siguiente condición de integrabilidad:

$$\int_{\mathbb{R}} p_t(x) |u_0| (dx) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi t}} \exp \{-x^2/2t\} |u_0| (dx),$$

para $t > 0$ y $x \in \mathbb{R}$. Estamos entonces interesados en el comportamiento asintótico de un nuevo objeto construido a partir de dicha solución: los promedios espaciales de la misma. Concretamente, en el caso en el que $u_0(x) = 1$, para cualquier $x \in \mathbb{R}$, y $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ es una función Lipschitz tal que $\sigma(1) \neq 0$, se definen los promedios espaciales de la solución u como sigue: para $R > 0$ fijo,

$$F_{R,t} := \frac{1}{\sigma_{R,t}} \left(\int_{-R}^R u(t, x) dx - 2R \right), \quad \text{donde } \sigma_{R,t}^2 = \text{Var} \left(\int_{-R}^R u(t, x) dx \right). \quad (4.2)$$

Las variables $F_{R,t}$ son conocidas como los promedios espaciales centrados y normalizados. En [3], Huang, Nualart y Viitasaari estudian el comportamiento límite de $F_{R,t}$ conforme $R \rightarrow \infty$, siendo capaces de probar resultados tipo teorema central del límite para $F_{R,t}$. Más aún, en dicha publicación, utilizando técnicas de Malliavin-Stein, se obtuvieron cotas superiores para la tasa de convergencia de $F_{R,t}$ a la distribución normal estándar en distancia de variación total. De manera precisa, obtuvieron que

$$d_{TV}(F_{R,t}, N) \leq \frac{C_t}{\sqrt{R}}, \quad \text{para } t > 0, \text{ y } R \geq 1. \quad (4.3)$$

Con este trabajo como antecedente, el propósito final del capítulo es obtener cotas superiores para la tasa de convergencia de la distancia uniforme entre las densidades de los promedios espaciales y la densidad de una variable aleatoria normal estándar. Concretamente, se busca probar el siguiente resultado.

Teorema 25. Sea $u = \{u(t, x) : (t, x) \in [0, \infty) \times \mathbb{R}\}$ la solución mild de la ecuación estocástica del calor dada antes. Supongamos que $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ es una función tal que

- $\sigma \in C^2(\mathbb{R})$.
- σ' es acotada.
- $|\sigma''(x)| \leq C(1 + |x|^m)$, para cualquier $x \in \mathbb{R}$ real y para alguna $m > 0$.

Supongamos también que, para algún $q > 10$, se cumple que

$$\mathbb{E} [|\sigma(u(t, 0))|^{-q}] < \infty.$$

Entonces para $t > 0$ fijo, se cumple que para cualquier $R > 0$, la distancia entre las densidades de los promedios espaciales y la densidad de la distribución normal están acotadas uniformemente en \mathbb{R} de la siguiente forma:

$$\sup_{x \in \mathbb{R}} |f_{F_{R,t}}(x) - \Phi(x)| \leq \frac{C_t}{\sqrt{R}},$$

donde $f_{F_{R,t}}$ y Φ son las densidades de $F_{R,t}$ y N una variable aleatoria normal estándar, respectivamente, y C_t es una constante dependiente de t .

La idea de la prueba es la siguiente. Dados los preliminares de SPDE's y cálculo de Malliavin-Stein, primero se demuestra una cota superior elemental para la distancia uniforme entre la densidad de una variable aleatoria F que está dada como un funcional de un proceso Gaussiano isonormal, y la densidad de una variable aleatoria normal estándar.

Para lograr dicha estimación, es necesario ver a la función de distribución de la variable aleatoria F como la integral de Skorokhod (o como divergencia en el sentido del cálculo de Malliavin). Esto es, $F = \delta(v)$ para algún v .

La idea es aplicar las técnicas de estimación anteriores a los promedios espaciales $F_{R,t}$ de la solución mild a la ecuación del calor.

Tenemos el siguiente resultado, el cual se encuentra en [1], y del cual haremos uso sin ahondar en su demostración.

Teorema 26. Sea $F \in \mathbb{D}^{1,1}$ y sea $v \in L^1(\Omega; \mathcal{H})$ tal que $D_v F \neq 0$ c.s. Supongamos que $v/D_v F \in \text{Dom}(\delta)$. Entonces la distribución de F tiene una densidad continua y acotada dada por

$$f_F(x) = \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \delta \left(\frac{v}{D_v F} \right) \right]$$

El resultado anterior es válido para cualquier espacio de Hilbert \mathfrak{H} . Aplicando el resultado anterior al contexto de una variable aleatoria adecuada, se tiene lo siguiente:

Teorema 27. Sea $v \in \mathbb{D}^{1,6}(\Omega; \mathcal{H})$ y $F = \delta(v) \in \mathbb{D}^{2,6}$ con $\mathbb{E}[F] = 0, \mathbb{E}[F^2] = 1$ y $(D_v F)^{-1} \in L^4(\Omega)$. Entonces $v/D_v F \in \text{Dom}(\delta)$, F admite una densidad $f_F(x)$ y se tiene la siguiente desigualdad

$$\sup_{x \in \mathbb{R}} |f_F(x) - \Phi(x)| \leq (\|F\|_4 \| (D_v F)^{-1} \|_4 + 2) \|1 - D_v F\|_2 + \| (D_v F)^{-1} \|^2 \|D_v (D_v F)\|_2,$$

donde $\Phi(x)$ es la función de densidad de una variable normal estándar.

Demostración. Vamos a denotar por F_R a los promedios espaciales de la ecuación estocástica del calor, y directamente por σ a la varianza de los promedios espaciales.

Buscamos hallar una cota uniforme para $x \in \mathbb{R}$ para las siguientes cantidades

$$|f_F(x) - \phi(x)|.$$

Gracias a que los promedios espaciales admite una densidad dada por

$$f_F(x) = \mathbb{E} \left[\mathbb{1}_{\{F_R > x\}} \delta \left(\frac{v}{D_v F_R} \right) \right],$$

podemos operar con estas cantidades de la siguiente forma:

$$\begin{aligned} f_F(x) &= \mathbb{E} \left[\mathbb{1}_{\{F_R > x\}} \delta \left(\frac{v}{D_v F_R} \right) \right] = \mathbb{E} \left[\mathbb{1}_{\{F_R > x\}} \frac{1}{D_{v_R}} \delta(v_R) \right] - \mathbb{E} \left[\mathbb{1}_{\{F_R > x\}} \langle D \left(\frac{1}{D_{v_R} F_R} \right), v_R \rangle \right] \\ &= \mathbb{E} \left[\mathbb{1}_{\{F_R > x\}} \frac{F_R}{D_{v_R} F_R} \right] + \mathbb{E} \left[\mathbb{1}_{\{F_R > x\}} \left(\frac{1}{(D_{v_R} F_R)^2} \right) D_{v_R} D_{v_R} F_R \right]. \end{aligned}$$

Ahora bien, en el término de la derecha de antes, podemos acotar la cantidad de la siguiente forma:

$$\mathbb{E} \left[\mathbb{1}_{\{F_R > x\}} \left(\frac{1}{(D_{v_R} F_R)^2} \right) D_{v_R} D_{v_R} F_R \right] \leq \sqrt{\mathbb{E} [|D_{v_R} F_R|^{-4}] \mathbb{E} [|D_{v_R} D_{v_R} F_R|^2]},$$

de tal forma que el problema simplificado consiste en analizar el término de la izquierda, es decir, analizar.

$$\mathbb{E} \left[\mathbb{1}_{\{F_R > x\}} \frac{F_R}{D_{v_R} F_R} \right]$$

Pero en cuanto a este término, notamos lo siguiente:

$$\mathbb{E} \left[\mathbb{1}_{[x, \infty)}(F_R) \frac{F_R}{D_v F_R} \right] = \mathbb{E} \left[\mathbb{1}_{[x, \infty)}(F_R) F_R \left(\frac{1}{D_v F_R} - \frac{1}{\sigma^2} \right) \right] + \frac{1}{\sigma^2} \mathbb{E} [\mathbb{1}_{[x, \infty)}(F_R) F_R]$$

Ahora bien, nosotros tenemos la siguiente estimación

$$\sigma^2 \approx \mathbb{E} [D_{V_R} F_R] \approx 1,$$

por lo que insertando dicha aproximación en la igualdad anterior, tenemos que

$$\begin{aligned} \mathbb{E} \left[\mathbb{1}_{[x, \infty)}(F_R) F_R \left(\frac{1}{D_v F_R} - \frac{1}{\sigma^2} \right) \right] + \frac{1}{\sigma^2} \mathbb{E} [\mathbb{1}_{[x, \infty)}(F_R) F_R] \\ = \frac{1}{(D_v F_R) \sigma^2} \mathbb{E} [\mathbb{1}_{[x, \infty)}(F_R) F_R] (\mathbb{E} [D_v F_R] - D_v F_R) \\ + \frac{1}{\sigma^2} \mathbb{E} [\mathbb{1}_{[x, \infty)}(F_R) F_R] \end{aligned}$$

Tenemos aquí dos términos nuevamente. El primero de ellos lo podemos acotar de la siguiente forma:

$$\|F_R\|_{L^4(\Omega)} \| (D_{v_R} F_R)^{-1} \|_{L^4(\Omega)} \text{Var} |D_{v_R} F_R|^{1/2} \leq C \|D D_{v_R} F_R\|_{L^2(\Omega, \langle \cdot \rangle)}^{1/2},$$

donde C es una constante adecuada.

Resta entonces analizar qué sucede con el siguiente término:

$$\mathbb{E} [\mathbb{1}_{[x, \infty)}(F_R) F_R].$$

Para ello, utilizamos la herramienta de Malliavin-Stein. Concretamente notemos que, sumando y restando el término $\mathbb{1}_{\{N > x\}} N$, donde N es una variable aleatoria normal estándar, tenemos que, denotando por $g_x(F_R) := \mathbb{1}_{\{F_R > x\}} F_R$ y lo correspondiente para $g_x(N)$,

$$\mathbb{E} [\mathbb{1}_{\{F_R > x\}}] = \mathbb{E} [g_x(F_R) - g_x(N)] + \mathbb{E} [g_x(N)].$$

El término de la izquierda en la ecuación anterior es sencillo de controlar, por lo que resta analizar la diferencia entre las esperanzas de g_x evaluada en F_R y en N . Para ello, de acuerdo al capítulo de método de Stein, notamos que

$$\mathbb{E} [g_x(F_R) - g_x(N)] = \mathbb{E} [F_R \psi_x(F_R) - \psi'_x(F_R)],$$

donde $y\psi_x(y) - \psi'_x(y) = g_x(y) - \mathbb{E}[g_x(N)]$.

Ahora bien, analizando la esperanza en términos de ψ_x , tenemos que

$$\begin{aligned}
\mathbb{E}[F_R\psi_x(F_R)] &= \mathbb{E}[\delta(v_R)\psi_x(F_R)] \\
&= \mathbb{E}[\langle v_R, DF_R \rangle | \psi'_x(F_R)] \\
&= \mathbb{E}[D_{V_R}F_R\psi'_x(F_R)] \\
&= \sigma^2\mathbb{E}[\psi_x(F_R)] + \mathbb{E}[(D_{v_R}F_R - \sigma^2)\psi'_x(F_R)] \\
&\leq \|\psi'_x\|_\infty \text{Var}[D_{v_R}F_R]^{-1/2} \\
&\leq \|\psi'_x\|_\infty \|DD_{v_R}F_R\|_{L^2(\Omega, \mathfrak{H})}.
\end{aligned}$$

Finalmente, para la prueba de la última cota, si nosotros vemos a la variable aleatoria normal N como una divergencia, esto es, $N = \delta(h)$ para algún h tal que $\|h\| = 1$, entonces

$$f_N(x) = \mathbb{E}[\mathbf{1}_{\{N > x\}}\delta(h)] = \mathbb{E}[\mathbf{1}_{\{N > x\}}N] = \mathbb{E}[g_x(N)],$$

por lo que concluimos la cota. ■

Y para la prueba del resultado principal, es esencial el uso de dos resultados distintos: uno de ellos consiste en una cotas de momentos para la segunda derivada de Malliavin de la solución, y los momentos negativos de la proyección $DF_{R,t}$ en $v_{R,t}$, donde $F_{R,t} = \delta(v_{R,t})$. El primero de ellos se encuentra en el siguiente resultado:

Teorema 28. Sea u la solución a la ecuación estocástica del calor con condición inicial $u_0 = 1$. Supongamos que las hipótesis del resultado principal. Fijando $(t, x) \in [0, \infty) \times \mathbb{R}$. Entonces $u(t, x) \in \bigcap_{p \geq 2} \mathbb{D}^{2,p}$ y para casi todo $0 < r < s < t$, $y, z \in \mathbb{R}$, la segunda derivada de u , a saber, $D_{r,z}D_{s,y}u(t, x)$ satisface la siguiente ecuación diferencial estocástica lineal:

$$\begin{aligned}
D_{r,z}D_{s,y}u(t, x) &= p_{t-s}(x - y)\sigma'(u(s, y))D_{r,z}u(s, y) \\
&\quad + \int_{[s,t] \times \mathbb{R}} p_{t-\tau}(x - \xi)\sigma''(u(\tau, \xi))D_{r,z}u(\tau, \xi)D_{s,y}u(\tau, \xi)W(d\tau, d\xi) \\
&\quad + \int_{[s,t] \times \mathbb{R}} p_{t-\tau}(x - \xi)\sigma'(u(\tau, \xi))D_{r,z}D_{s,y}u(\tau, \xi)W(d\tau, d\xi)
\end{aligned}$$

Más aún, para cualesquiera $0 \leq r < s < t \leq T$ y $x, y, z \in \mathbb{R}$, tenemos que

$$\|D_{r,z}D_{s,y}u(t, x)\|_p \leq C_{T,p}\Phi_{r,z,s,y}(t, x),$$

donde $C_{T,p}$ es una constante que depende de T, p y σ , y

$$\Phi_{r,z,s,y}(t,x) := p_{t-s}(x-y) \cdot \left(p_{s-r}(y-z) + \frac{p_{t-r}(z-y) + p_{t-r}(z-x) + \mathbf{1}_{|y-x|>|z-y|}}{(s-r)^{1/4}} \right)$$

Por otro lado, se tiene el siguiente resultado sobre la estimación de momentos negativos.

Teorema 29. Sea u la solución mild a la ecuación estocástica del calor. Supongamos que σ es Lipschitz. Sea $p \geq 2$ fijo, $t > 0$ y supongamos que existe un $q > 5p$ tal que $\mathbb{E} [|\sigma(u(t,0))|^{-2q}] < \infty$. entonces, existe una constante $R_0 > 0$ tal que

$$\sup_{R \geq R_0} \mathbb{E} \left[|D_{v_{R,t}} F_{R,t}|^{-p} \right] < \infty.$$

La prueba de este resultado se omite en este texto, por lo que se refiere a [5] para su consulta.

4.1. Formulación Mild en el caso sin deriva

4.2. Densidad explícita de los promedios espaciales

Contribución del Kernel de Stein Contribución de las derivadas de orden superior

4.3. Aproximación uniforme de las densidades

Bibliografía

- [1] María Emilia Caballero, Begoña Fernández, and David Nualart, *Estimation of densities and applications*, Journal of Theoretical Probability **11** (1998), no. 3, 831–851.
- [2] Yaozhong Hu, Fei Lu, and David Nualart, *Convergence of densities of some functionals of gaussian processes*, Journal of Functional Analysis **266** (2014), no. 2, 814–875.
- [3] Jingyu Huang, David Nualart, and Lauri Viitasaari, *A central limit theorem for the stochastic heat equation*, Stochastic Processes and their Applications **130** (2020), no. 12, 7170–7184.
- [4] Davar Khoshnevisan, *A primer on stochastic partial differential equations*, pp. 1–38, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [5] Sefika Kuzgun and David Nualart, *Convergence of densities of spatial averages of stochastic heat equation*, Stochastic Processes and their Applications **151** (2022), 68–100.
- [6] Jean-François Le Gall, *Brownian motion, martingales, and stochastic calculus*, Graduate Texts in Mathematics, Springer International Publishing, 2016.
- [7] Ivan Nourdin and Giovanni Peccati, *Normal approximations with malliavin calculus: From stein's method to universality*, Cambridge Tracts in Mathematics, Cambridge University Press, 2012.
- [8] John B. Walsh, *An introduction to stochastic partial differential equations*, École d'Été de Probabilités de Saint Flour XIV - 1984 (Berlin, Heidelberg) (P. L. Hennequin, ed.), Springer Berlin Heidelberg, 1986, pp. 265–439.