

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

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

En memoria de mi madre.

Agradecimientos

A mis padres . . .

Índice general

Agradecimientos	III
Introducción	VII
1. Elementos de ecuaciones diferenciales parciales estocásticas	1
1.1. Introducción a las SPDE	2
1.1.1. Soluciones débiles de una EDP	2
1.1.2. 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	13
1.2.3. Definición de la integral y algunas propiedades	18
1.3. Formulación mild de una SPDE. Existencia y unicidad de las soluciones	24
2. Elementos de Cálculo de Malliavin	29
2.1. Cálculo de Malliavin en el caso unidimensional	29
2.1.1. La derivada de Malliavin.	29
2.1.2. El operador de divergencia	36
2.1.3. El semigrupo de Ornstein-Uhlenbeck	37
2.1.4. El generador infinitesimal	43
2.2. Cálculo de Malliavin en el caso general.	46
3. Aplicaciones a Teoremas Límite y estudio de densidades	51
3.1. Método de Stein	51
3.2. Existencia de densidades y estimación de la distancia uniforme	58
4. Propiedades espaciales de la Ecuación estocástica del calor	63
4.1. Cotas para la primera y segunda derivada de la solución mild y existencia de momentos inversos	65
4.2. Prueba del resultado principal	66
5. Conclusiones y trabajo futuro	83

A. Resultados técnicos

87

Introducción

Este trabajo está basado principalmente en el artículo de Şefika Kuzgun y David Nualart [11] publicado en el año 2022. El objetivo principal es estudiar la convergencia de las densidades de los promedios espaciales de la solución de la ecuación del calor estocástica (SHE, por sus siglas en inglés), en distancia uniforme, a la densidad de una distribución normal estándar.

La presente tesis aborda la temática que rodea a dicho problema con base en dos piezas fundamentales. La primera consiste en describir las densidades de variables aleatorias en términos de operadores fundamentales del Cálculo de Malliavin. Trabajo previo con respecto al estudio de existencia de densidades se puede hallar desde el artículo de 1998 de Maria Emilia Caballero, Begoña Fernández Fernández y David Nualart [2], en donde se demuestra la existencia y unicidad de una densidad continua y acotada para una variable aleatoria X que cumpla ciertas condiciones. La segunda pieza consiste en implementar el método de Malliavin-Stein para estudiar la cuantificación del error en aproximaciones gaussianas. En esta dirección, técnicas que mezclan al cálculo de Malliavin junto con el método de Stein son propuestas por Nourdin y Peccati en su monografía [14] publicada en el año 2012. Dicha monografía muestra que el denominado método de Malliavin-Stein resulta muy útil a la hora de establecer teoremas tipo límite central y obtener tasas de convergencia explícitas para los mismos.

Entre los antecedentes del artículo de nuestro interés [11] podemos citar el trabajo de Yaozhong Hu, Fei Lu y David Nualart [8]. Allí se estudia la convergencia de las densidades de algunos procesos gaussianos. Posteriormente, en 2020, Jingyu Huang, Lauri Viitasaari y David Nualart en [9] logran demostrar un teorema central del límite para la solución a la ecuación del calor estocástica. Más aún, ellos demuestran cotas explícitas para la velocidad de convergencia de la ley de la solución a la ecuación del calor estocástica a la distribución normal, medida en términos de la distancia en variación total.

Con base en dichos antecedentes, Şefika Kuzgun y David Nualart demuestran la existencia y unicidad de las densidades de los promedios espaciales de la solución a la ecuación del calor estocástica, y demuestran la convergencia de estas densidades, en distancia uniforme, a la densidad de una variable aleatoria normal, otorgando una tasa de convergencia explícita para la misma.

Es de destacar que la intención de este texto no es la de dar un estudio exhaustivo de las teorías

que forman parte de la herramienta para abordar el problema. Lo que se busca es cubrir de manera suficiente pero lo más sucinto posible estas herramientas técnicas con el objetivo de abordar el problema. Por lo tanto, en esta tesis se da un breve recorrido por las ideas que yacen en el núcleo de la teoría de ecuaciones diferenciales parciales estocásticas (SPDE, por sus siglas en inglés), del cálculo de Malliavin, del método de Stein, así como de la combinación que de estas últimas dos disciplinas surge con la intención de estudiar la existencia de densidades de variables aleatorias bajo ciertas hipótesis, y la deducción de teoremas límite.

Lo anterior motiva la estructura del presente texto. En el primer capítulo se estudiarán los elementos esenciales de ecuaciones diferenciales parciales estocásticas. Específicamente, se enuncia brevemente la heurística que da origen a la formulación de las mismas. Se realiza la construcción de la integral de Itô-Walsh, y con dicha integral se formula la solución mild de una ecuación diferencial parcial estocástica. Posteriormente nos enfocamos en el contexto de la ecuación del calor estocástica, así como en su solución. Finalmente se enuncian algunos resultados técnicos necesarios para abordar el problema principal, destacando las desigualdades de Burkholder-Davis-Gundy y una versión adecuada del teorema de Fubini.

El capítulo dos está dedicado al estudio de los elementos básicos del cálculo de Malliavin. Este capítulo se divide a su vez en dos partes. En la primera se hace un estudio del caso unidimensional, siendo este caso estudiado con mediano detalle, aunque siempre manteniéndose al mínimo la teoría requerida para nuestro propósito. La segunda parte del capítulo se compone del caso general. Aquí nos limitamos principalmente a enunciar los resultados análogos al caso unidimensional.

La composición del capítulo tres es similar. La primera parte del mismo consiste en un estudio breve del método de Stein, mencionando tangencialmente algunas aplicaciones de esta disciplina. La segunda parte del capítulo consiste principalmente en estudiar un resultado que garantiza la existencia y unicidad de la densidad de una variable aleatoria, así como en estudiar la existencia de cotas para la distancia uniforme entre la densidad de una variable adecuada y la densidad de una variable normal estándar. Este último resultado es la columna sobre la que se sostiene el teorema principal, y es demostrado con detalle. Además es en esta parte en donde se hace uso de la combinación entre cálculo de Malliavin y el método de Stein. El estudio de las herramientas necesarias para abordar el problema principal culmina en esta parte.

El enunciado y prueba del resultado principal se encuentran hasta el capítulo cuarto. Es aquí en donde se hace rigurosa la formulación del problema que se plantea. Se enuncian dos lemas de carácter técnico relacionados con momentos inversos y estimaciones de varianzas para las derivadas de Malliavin de los promedios espaciales para la ecuación del calor estocástica. Estos lemas son esenciales en la prueba. Finalmente se estudia y demuestra el resultado principal. Se agrega un último capítulo en el cual se hace una breve reflexión sobre lo realizado, y se echa una mirada al panorama en el cual este problema está inmerso, enunciando posibles direcciones para trabajo futuro.

Los primeros tres capítulos comienzan con un muy breve resumen histórico de la disciplina a estudiar, así como de la relevancia de la misma en la actualidad. El cuarto capítulo repasa el trabajo previo que ha sido mencionado al inicio de esta introducción, pero con mayor rigurosidad.

La intención del autor principalmente radica en explicar de la manera más simple que le ha sido posible los resultados, así como rellenar algunos detalles en las pruebas, o bien unificar algunas ideas sueltas de las fuentes consultadas.

Palabras clave: *Integral de Itô-Walsh, Ecuación del calor estocástica, promedios espaciales, cálculo de Malliavin, método de Stein, existencia de densidades, convergencia uniforme, teoremas límite, tasas de convergencia.*

Capítulo 1

Elementos de ecuaciones diferenciales parciales estocásticas

En este capítulo hablaremos acerca de ecuaciones diferenciales parciales estocásticas o SPDEs. 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 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 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 [18], escrita por Walsh y publicada en 1986, en donde el autor aborda el estudio de estas 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 tanto una variable espacial 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 un área bastante activa en la probabilidad moderna.

En la primera parte de este capítulo, hacemos un resumen de las soluciones débiles, a una ecuación diferencial parcial clásica. Dichas ideas son clave para estudiar el caso estocástico. La idea de la integral de Itô-Walsh surge de manera natural en esta parte. Es importante destacar que esta primera sección introductoria se presenta solo ideas y razonamientos formales, dejando la parte rigurosa y en el contexto que nos interesa para el final del capítulo. Dichas ideas están basadas en [6], [7] y [10].

1.1. Introducción a las SPDE

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 del calor estocástica homogénea clásica, la cual está dada por

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

una solución clásica es una función $u : \mathbb{R}^N \times [0, \infty) \rightarrow \mathbb{R}$ tal que $u \in C^{2,1}(\mathbb{R}^N \times [0, \infty))$ (una función dos veces continuamente derivable en las variables espaciales, y continuamente diferenciable en la variable temporal), que cumpla la ecuación anterior. En particular, la solución clásica cumple que al menos todas las segundas derivadas en las variables espaciales existen, 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.

1.1.1. Soluciones débiles de una EDP

Con motivo de resolver problemas como el planteado anteriormente, pero en los que en principio la regularidad es un problema, surge la idea de las *soluciones débiles* de las ecuaciones diferenciales parciales. Un ejemplo de ellos es aquél problema de onda con una condición inicial que no necesariamente es regular. Por ejemplo, si hablamos del siguiente problema de onda

unidimensional homogéneo 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 $\pi - |x|$ pues 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 cero en cualquier punto de la cuerda.

La idea para hallar una solución al problema anterior, a grandes rasgos, consiste en reformularlo de forma que se conserve la esencia del mismo, 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 $\phi \in C^\infty([-\pi, \pi] \times [0, \infty))$ tal que sus derivadas se anulan en la frontera de $[-\pi, \pi]$, se tiene que al multiplicar e integrar, la ecuación anterior se convierte en:

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

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

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

en donde los términos evaluados en la frontera se anulan gracias a que las funciones ϕ (a menudo denominadas funciones de prueba o funciones *test*) 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, vamos a decir que u es una *solución débil* de la ecuación diferencial asociada al problema (1.2) si se cumple que

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

para cualquier función $\phi \in C^\infty([-\pi, \pi] \times [0, \infty))$ cuyas derivadas se anulen 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 gozan 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, busquemos incluir ideas probabilistas. 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 (ecuación que será de nuestro interés),

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

en donde el término F es aleatorio, ¿qué significa hablar de una solución a 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 el problema (1.2) en un sentido débil, nos encontramos con una dificultad pues $F(x, t)$ es ahora un proceso estocástico, y a no ser que dicho proceso sea de variación finita, no será posible realizar de manera rigurosa la transición al momento de integrar por partes.

1.1.2. Heurística de la formulación mild

Apoyándonos en las ideas de las soluciones débiles, un momento de reflexión puede arrojar luz sobre cómo replantear de manera rigurosa un *problema del calor estocástico* como el anterior. Haciendo nuevamente cálculos formales, una manera de proceder sería la siguiente: supongamos que tenemos la siguiente ecuación diferencial del calor unidimensional no lineal,

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

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 solución débil a la ecuación de onda, podríamos tomar una función $\phi \in C^\infty([0, L])$ cuya derivada se anule en la frontera, multiplicarla en ambos miembros de la ecuación anterior e integrar con respecto al espacio y al tiempo, y suponiendo que de alguna manera logramos integrar con respecto la fuente aleatoria $W(t, x)$, tendremos la siguiente ecuación formal,

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

Es importante destacar que hasta ahora lo que significa integrar con respecto a la fuente aleatoria W es desconocido, pero una vez que la integral anterior haya cobrado sentido, definir con rigor lo que significa un problema como el planteado anteriormente y lo que significa resolverlo es similar a la idea de las soluciones débiles, aunque por supuesto que con mayor complejidad debido a la naturaleza estocástica del problema. Dichas soluciones se conocen en probabilidad con el nombre de *soluciones mild*.

Echando mano de la intuición que la integral de Itô nos da al formalizar la idea de integrar una función con respecto a un movimiento browniano, nos resta estudiar una manera de dar rigor a la integral anterior, al menos suponiendo que la fuente de aleatoriedad W de dos variables proviene de distribuciones gaussianas, las cuales son relativamente «nobles» en su manejo.

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. Esta integral resuelve el problema planteado anteriormente sobre cómo integrar con respecto a una fuente de aleatoriedad W . 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 [13] 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, X tiene distribución normal estándar si 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 $X \sim \text{Normal}(0,1)$.

Directamente de la definición, tenemos la caracterización de una 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 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 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 Y tiene media $\mu \in \mathbb{R}$ y varianza $\sigma^2 > 0$ si se cumple que $Y = \sigma X + \mu$, donde X es una variable normal estándar. Denotaremos la situación anterior escribiendo $Y \sim \text{Normal}(\mu, \sigma^2)$. En vista de la proposición 1 y de la definición, se tiene el siguiente resultado.

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.

- $\mathcal{L}_Y(\lambda) = e^{\mu\lambda + \sigma^2\lambda^2/2}, \quad \lambda \in \mathbb{C},$
- $\mathcal{F}_Y(\xi) = e^{i\mu\xi - \sigma^2\xi^2/2}, \quad \xi \in \mathbb{R},$

- 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 $\sigma^2 = 0$ si se cumple que $Y = \mu$, salvo un conjunto de medida de Lebesgue cero. Pasamos ahora a definir los vectores aleatorios 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 $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ 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 un vector $\mu_X \in \mathbb{R}^d$ 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 podemos caracterizar a los objetos anteriores, pues 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 a $u \in \mathbb{R}^d$ como $u = \sum_{j=1}^d u_j e_j$,

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

Se deduce 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

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

En vista de las dos proposiciones anteriores, tenemos 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 $G(t)$, con $t \in T$, 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 vectores aleatorios gaussianos.

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

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, los siguientes hechos son equivalentes.

- Los subespacios H_1 y H_2 son ortogonales entre sí en $L^2(\mathbb{P})$.
- Las σ -álgebras $\sigma(G_1)$ y $\sigma(G_2)$ son independientes.

Es importante recordar que esta última propiedad es muy particular del contexto gaussiano, pues en conjuntos de variables aleatorias que no sigan posean dicha ley, no necesariamente es equivalente la ortogonalidad entre los subespacios generados por las variables aleatorias, a la independencia de las σ -álgebras generadas por las mismas. 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 medias y covarianzas 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} determinan 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 , se expresa en términos de la función Γ definida antes mediante la fórmula $\Sigma_X = (\text{Cov}(G(t_i), G(t_j)))_{1 \leq i, j \leq k} = (\Gamma(t_i, t_j))_{1 \leq i, j \leq k}$.

El hecho de que la función de covarianzas Γ sea no negativa definida es consecuencia de que la forma cuadrática q_X para cualquier vector X que componga una distribución finito-dimensional de G , es no negativa definida y simétrica. De manera más precisa, 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 conocerlas para reconstruir el proceso gaussiano. Éste es el contenido del siguiente teorema.

Proposición 7. Sean $\Gamma : T \times T \rightarrow \mathbb{R}$ y $\mu : T \rightarrow \mathbb{R}$ funciones 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 medias y covarianzas del proceso G .

La demostración de lo anterior es una consecuencia del Teorema de Consistencia (o extensión) de Kolmogorov (ver Teorema 6.3 en [13]). 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 para $c : T \rightarrow \mathbb{R}$ 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 7, existe en un espacio de probabilidad adecuado un proceso gaussiano que denotaremos por $(B(t))_{t \geq 0}$, con función de medias la constante 0 y función de covarianzas $\Gamma(s, t) = s \wedge t$. Dicho proceso es el *movimiento Browniano estándar* en \mathbb{R} .

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 definimos $\Gamma(A, B) := \lambda^d(A \cap B)$, donde $\lambda^d : \mathcal{B}(\mathbb{R}^d) \rightarrow [0, \infty)$ 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, éstas 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, denotando por $V = \text{Var}(\dot{W}(A \cup B) - \dot{W}(A) - \dot{W}(B) + \dot{W}(A \cap B))$, y usando las propiedades de la covarianza,

$$\begin{aligned} V &= \text{Cov}(\dot{W}(A \cup B), \dot{W}(A \cup B)) + \text{Cov}(\dot{W}(A), \dot{W}(A)) + \text{Cov}(\dot{W}(B), \dot{W}(B)) \\ &\quad + \text{Cov}(\dot{W}(A \cap B), \dot{W}(A \cap B)) - 2\text{Cov}(\dot{W}(A \cup B), \dot{W}(A)) - 2\text{Cov}(\dot{W}(A \cup B), \dot{W}(B)) \\ &\quad + 2\text{Cov}(\dot{W}(A \cup B), \dot{W}(A \cap B)) + 2\text{Cov}(\dot{W}(A), \dot{W}(B)) - 2\text{Cov}(\dot{W}(A), \dot{W}(A \cap B)) \\ &\quad - 2\text{Cov}(\dot{W}(B), \dot{W}(A \cap B)). \end{aligned}$$

Usando ahora la definición de Γ , la expresión anterior nos dice que V cumple que

$$\begin{aligned}
V &= \lambda^d(A \cup B) + \lambda^d(A) + \lambda^d(B) + \lambda^d(A \cap B) - 2\lambda^d((A \cup B) \cap A) - 2\lambda^d((A \cup B) \cap B) \\
&\quad + 2\lambda^d((A \cup B) \cap (A \cap B)) + 2\lambda^d(A \cap B) - 2\lambda^d(A \cap (A \cap B)) - 2\lambda^d(B \cap (A \cap B)) \\
&= \lambda^d(A \cup B) + \lambda^d(A) + \lambda^d(B) + \lambda^d(A \cap B) - 2\lambda^d(A) - 2\lambda^d(B) + 2\lambda^d(A \cap B) \\
&\quad + 2\lambda^d(A \cap B) - 2\lambda^d(A \cap B) - 2\lambda^d(A \cap B) \\
&= \lambda^d(A \cup B) - \lambda^d(A) - \lambda^d(B) + \lambda^d(A \cap B)
\end{aligned}$$

De aquí se sigue que $V = 0$. Por lo tanto, $\dot{W}(A \cup B) - (\dot{W}(A) + \dot{W}(B) - \dot{W}(A \cap B))$ 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 [10]). 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. Utilizamos ahora la bien conocida heurística de teoría de la medida para definir nuevos objetos a partir de otros más sencillos, y luego extender por aproximación. Siguiendo esta maquinaria estándar, para un conjunto $A \in \mathcal{B}(\mathbb{R}^d)$ definimos $\dot{W}(\mathbf{1}_A) := \dot{W}(A)$, 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

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

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, al ser centradas,

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

y usando nuevamente la definición de Γ ,

$$\begin{aligned} &= \sum_{k=1}^n c_k^2 \Gamma(A_k, A_k) \\ &= \sum_{k=1}^n c_k^2 \lambda^d(A_k) \\ &= \int_{\mathbb{R}^d} \sum_{k=1}^n c_k^2 \mathbb{1}_{A_k}(x) \lambda^d(dx) \\ &= \left\| \sum_{k=1}^n c_k \mathbb{1}_{A_k}(x) \right\|_{L^2(\mathbb{R}^d)}^2, \end{aligned}$$

donde en la última igualdad utilizamos que $A_i \cap A_j = \emptyset$ y por lo tanto $\mathbb{1}_{A_i} \mathbb{1}_{A_j} = 0$ 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 también 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, 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 $f \mapsto \dot{W}(f)$ 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

$$\dot{W}(f) = \int_{\mathbb{R}^d} f(x) W(dx).$$

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

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

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

Como veremos más adelante, los procesos isonormales juegan un papel importante en el Cálculo de Malliavin.

1.2.2. Medidas martingala

En esta parte del texto, estudiamos el concepto de medidas martingala. Dichos procesos estocásticos son una piedra 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.

Teorema 1. El proceso $(\dot{W}(A))_{A \in \mathbb{R}^d}$ cumple las siguientes propiedades

- $\dot{W}(\emptyset) = 0$, \mathbb{P} – 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, pero con valores en $L^2(\mathbb{P})$.

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 (igualdad (1.5)), 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 notar que el ruido blanco forma una medida con valores en $L^2(\mathbb{P})$, pero esto es distinto a ser una medida signada aleatoria en los borelianos de \mathbb{R}^d . La diferencia es sutil, y radica la segunda propiedad de la proposición anterior, pues el cuantificador se encuentra 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

$$\mathcal{F}_t := \sigma \left(\dot{W}([0, t] \times A) : 0 \leq s \leq t, A \in \mathcal{B}(\mathbb{R}^d) \right), \quad t \geq 0. \quad (1.6)$$

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 signada, con valores en $L^2(\mathbb{P})$, y también una martingala. De manera precisa, se tiene el siguiente resultado.

Teorema 2. Sea $(W_t(A))_{t \geq 0, A \in \mathcal{B}(\mathbb{R}^d)}$ un proceso de ruido blanco en $[0, \infty) \times \mathbb{R}^d$. Entonces dicho proceso cumple las siguientes propiedades.

- Para cualquier $A \in \mathcal{B}(\mathbb{R}^d)$, se cumple la igualdad $W_0(A) = 0$, \mathbb{P} -casi seguramente.
- Si $t > 0$, entonces la aplicación $A \mapsto W_t(A)$ es una medida signada, σ -finita, con valores en $L^2(\mathbb{P})$.
- Para cualquier $A \in \mathbb{R}^d$, el proceso $(W_t(A))_{t \geq 0}$ 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 la aplicación $W_t : \mathcal{B}(\mathbb{R}^d) \rightarrow L^2(\mathbb{P})$ sea una medida signada σ -finita se sigue de la propiedad siguiente: para cualquier $A \in \mathcal{B}(\mathbb{R}^d)$,

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

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 1: 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 $(W_t(A))_{t \geq 0}$ es una martingala con respecto a la filtración $(\mathcal{F}_t)_{t \geq 0}$, observamos que el proceso claramente es integrable, pues 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 $W_{t+s} - W_t$ 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 un proceso que funciona, en cierto sentido, tanto como una medida como una martingala. Este concepto se puede definir de manera rigurosa, 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, y con valores en $L^2(\mathbb{P})$.
- Para cualquier $A \in \mathcal{B}(\mathbb{R}^d)$, se tiene que $(M_t(A))_{t \geq 0}$ es una martingala de media cero con respecto a la filtración $(\mathcal{F}_t)_{t \geq 0}$.

Por construcción, el proceso del ruido blanco es un primer ejemplo de medida martingala. El motivo para la introducción de este concepto es que estos objetos forman una clase adecuada de procesos que pueden fungir como integradores. No obstante, no cualquier medida martingala resulta ser un integrador adecuado. A partir de este punto de la sección, buscaremos establecer las propiedades que una medida martingala debe poseer para poder fungir como un integrador.

A continuación, definimos un concepto relacionado íntimamente con las medidas martingala, y que se puede interpretar como una generalización del proceso de covariación para dos martingalas en el caso unidimensional.

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

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

en donde se usa la notación de bracket usada para el proceso de covariación cuadrática entre dos procesos.

Gracias a que el funcional de covariación de una medida martingala está en términos de procesos de covariación, se deducen directamente de las propiedades de la covariación entre dos martingalas las siguientes propiedades del funcional de covariación.

Teorema 3. Sea M una medida martingala y denotemos por \overline{Q} a su funcional de covariación. 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,
- $t \mapsto \overline{Q}_t(A, A)$ es \mathbb{P} - casi seguramente no decreciente.

Con el concepto de funcional de covariación, podemos definir una función aleatoria de conjuntos, que denotaremos Q , como sigue:

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

Extendemos dicha función a uniones ajenas de elementos (rectángulos) $A_i \times B_i \times (s_i, t_i]$ donde $A_i, B_i \in \mathcal{B}(\mathbb{R}^d)$, $(s_i, t_i] \subseteq \mathbb{R}$ y $1 \leq i \leq m$, como

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

Como es estándar en teoría de la medida, se puede demostrar que la definición de dicha función Q no se ve alterada por un cambio en los representantes de los conjuntos. Siendo también un procedimiento rutinario, lo natural es pensar en extender la función aleatoria Q a una medida en $\mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(\mathbb{R}^d) \times \mathcal{B}([0, \infty))$, pero lamentablemente no es posible hacerlo para cualquier medida martingala. Es posible encontrar un contraejemplo en [18, pp. 305 - 307].

No obstante, sí es posible encontrar una clase de medidas martingala que sean útiles para nuestros propósitos. Para ello, estudiamos el concepto de *worthiness* de una medida martingala. 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.

Para ello, primero definimos lo que significa que una medida signada sea positiva definida en nuestro contexto.

Definición 8. Sea K una medida signada definida en la σ -álgebra $\mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(\mathbb{R}^d) \times \mathcal{B}([0, \infty))$. Decimos que K es positiva definida si para cualquier función f integrable y acotada, se tiene que

$$\int_{\mathbb{R}^d \times \mathbb{R}^d \times [0, \infty)} f(x, s) f(y, s) K(dx, dy, ds) \geq 0.$$

Introducimos ahora el concepto de *worthiness* de una medida martingala.

Definición 9. Sea M una medida martingala. Decimos que M es *worthy* (traducido literalmente al español como *digna*) si existe una medida σ -finita $K : \mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(\mathbb{R}^d) \times \mathcal{B}([0, \infty)) \times \Omega \rightarrow [0, \infty)$ tal que para cualesquiera $A, B \in \mathcal{B}(\mathbb{R}^d)$, $C \in \mathcal{B}([0, \infty))$ y $\omega \in \Omega$,

- La aplicación $A \times B \mapsto K(A \times B \times C, \omega)$ es positiva definida y simétrica, para C y ω fijos.
- $\{K(A \times B \times (0, t])\}_{t \geq 0}$ es un proceso predecible, para A y B fijos.
- 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 $t > 0$ y $A, B \in \mathcal{B}(\mathbb{R}^d)$,

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

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 . Como convención, utilizaremos el adjetivo *digna* para una medida martingala M que cumpla la definición anterior.

Lo relevante para nosotros es que, si M es una medida martingala digna y Q_M es la función aleatoria de conjuntos construida a partir de ella, entonces Q_M puede ser extendida como una medida signada a toda la σ -álgebra $\mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(\mathbb{R}^d) \times \mathcal{B}([0, \infty))$. Más aún, Q_M resulta ser positiva definida (ver [18, p. 291]).

1.2.3. Definición de la integral y algunas propiedades

Hemos reunido suficiente teoría para definir la integral de Itô-Walsh con respecto a medidas martingala. Conforme vayamos desarrollando la construcción, una cierta familiaridad con la construcción de la integral de Lebesgue y la integral de Itô será notoria.

Definición 10. Sean $(\mathcal{F}_t)_{t \geq 0}$ una filtración continua por derecha y M una medida martingala asociada a dicha filtración. Una función $f : \mathbb{R}^d \times [0, \infty) \times \Omega \rightarrow \mathbb{R}$ se dice que es *elemental* si es de la forma

$$f(x, t, \omega) = X(\omega) \mathbf{1}_{(a, b]}(t) \mathbf{1}_A(x), \quad (1.7)$$

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 la función elemental f con respecto a la medida martingala M como

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

Observamos en la definición una similitud con la construcción de Itô de una función (o bien, puede decirse, de un proceso) elemental con respecto a una martingala adecuada, pues en tal contexto, la variable X está multiplicada por un incremento de dicha martingala en el intervalo $(a, b]$. En el contexto de la integral de Itô-Walsh, la martingala encuentra su análogo en las medidas martingala M . Como una primera propiedad, tenemos la siguiente proposición.

Teorema 4. Sea f una función elemental como en (1.7) 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. Claramente, $(f \cdot M)_0(B) = 0$ por definición de la integral, para cualquier $B \in \mathcal{B}(\mathbb{R}^d)$. Por otro lado, $(f \cdot M)_t(\emptyset) = 0$ usando que M es medida martingala y por lo tanto es una medida signada cuando $t > 0$ es fijo. Ahora bien, si tenemos una unión numerable de conjuntos ajenos,

$$\begin{aligned} (f \cdot M)_t \left(\bigcup_{k \geq 1} B_k \right) &= X \left[M_{t \wedge b} \left(A \cap \bigcup_{k \geq 1} B_k \right) - M_{t \wedge a} \left(A \cap \bigcup_{k \geq 1} B_k \right) \right] \\ &= X \left[\sum_{k \geq 1} M_{t \wedge b}(A \cap B_k) - \sum_{k \geq 1} M_{t \wedge a}(A \cap B_k) \right] \\ &= \sum_{k \geq 1} (f \cdot M)_t(B_k), \end{aligned}$$

en donde hemos hecho uso nuevamente de que para $t > 0$ fijo, M se comporta como una medida signada. La σ -finitud de la integral se deduce de manera similar, y claramente se trata de una medida $L^2(\mathbb{P})$ -valuada.

Por último, fijando un conjunto $B \in \mathcal{B}(\mathbb{R}^d)$, tenemos que para $t \in (a, b]$ (el único caso en que f no se anula),

$$\begin{aligned}\mathbb{E}[(f \cdot M)_{t+s}(B)|\mathcal{F}_t] &= \mathbb{E}[X[M_{(t+s)\wedge b}(A \cap B) - M_{(t+s)\wedge a}(A \cap B)]|\mathcal{F}_t] \\ &= X\mathbb{E}[M_{(t+s)\wedge b}(A \cap B) - M_a(A \cap B)|\mathcal{F}_t] \\ &= X(\mathbb{E}[M_{(t+s)\wedge b}(A \cap B)|\mathcal{F}_t] - M_a(A \cap B)) \\ &= X(M_{t\wedge b} - M_{t\wedge a}) \\ &= (f \cdot M)_t(B),\end{aligned}$$

en donde hemos hecho uso de que X es \mathcal{F}_a medible y para B fijo, M se comporta como una martingala. ■

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 de la integral a funciones simples.

Definición 11. 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 8. La definición anterior es consistente, esto es, no depende de la representación de f en términos de funciones elementales.

También es posible demostrar que si $f \in \mathcal{S}$, entonces $(f \cdot M)$ es nuevamente una medida martingala.

Resulta ser que una clase adecuada de funciones integrables son las funciones f que son predecibles. Esto es, funciones medibles con respecto a aquella σ -álgebra que es generada por todas las funciones simples. Dicha σ -álgebra la denotamos por $\mathcal{P} := \sigma(\mathcal{S})$.

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

Teorema 5. Sea $f \in \mathcal{S}$ y supongamos que M es una medida martingala digna. Sea Q_M la función aleatoria de conjuntos asociada a M . Entonces para $B \in \mathcal{B}(\mathbb{R}^d)$,

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

Demostración. Primero hacemos la prueba para funciones elementales. Supongamos que f es de la forma (1.7). Entonces

$$\begin{aligned} \mathbb{E} [(f \cdot M)_t^2(B)] &= \mathbb{E} [X^2 (M_{t \wedge b}(A \cap B) - M_{t \wedge a}(A \cap B))^2] \\ &= \mathbb{E} [X^2 M_{t \wedge b}^2(A \cap B)] - 2\mathbb{E} [X^2 M_{t \wedge b}(A \cap B) M_{t \wedge a}(A \cap B)] \\ &\quad + \mathbb{E} [X^2 M_{t \wedge a}^2(A \cap B)]. \end{aligned}$$

Ahora bien, como M es medida martingala, es una martingala para $A \cap B$. Por lo tanto, por definición del proceso covariación, $M_t^2(A \cap B) - \langle M(A \cap B), M(A \cap B) \rangle_t$ es una martingala. Se sigue que

$$\mathbb{E} [M_{t \wedge b}^2(A \cap B) - \langle M(A \cap B), M(A \cap B) \rangle_{t \wedge b}] = \mathbb{E} [M_{t \wedge a}^2(A \cap B) - \langle M(A \cap B), M(A \cap B) \rangle_{t \wedge a}],$$

y al ser X una variable \mathcal{F}_a medible, también se cumple

$$\begin{aligned} &\mathbb{E} [X^2 (M_{t \wedge b}^2(A \cap B) - \langle M(A \cap B), M(A \cap B) \rangle_{t \wedge b})] \\ &= \mathbb{E} [X^2 (M_{t \wedge a}^2(A \cap B) - \langle M(A \cap B), M(A \cap B) \rangle_{t \wedge a})]. \end{aligned}$$

De manera similar, se tiene la igualdad

$$\begin{aligned} &\mathbb{E} [X^2 (M_{t \wedge b}^2(A \cap B) M_{t \wedge a}(A \cap B) - \langle M(A \cap B), M(A \cap B) \rangle_{t \wedge b})] \\ &= \mathbb{E} [X^2 (M_{t \wedge a}^2(A \cap B) M_{t \wedge a}(A \cap B) - \langle M(A \cap B), M(A \cap B) \rangle_{t \wedge a})]. \end{aligned}$$

Combinando estas expresiones, y recordando la definición de Q_M , obtenemos el resultado. La extensión al caso en que f es simple es rutinaria. ■

La reelevancia del resultado anterior se debe nuevamente a la analogía que existe con el cálculo estocástico convencional. En tal contexto, la isometría de Itô nos dice que la esperanza del cuadrado de la integral estocástica equivale a la esperanza de la integral, con respecto al proceso de covariación, del cuadrado del proceso en cuestión. En nuestro contexto, integrar con respecto al proceso de covariación del proceso tiene su análogo en integrar con respecto a la función de conjuntos Q , la cual vimos que forma una medida cuando la medida martingala M es digna, y está construida con base en los incrementos de los procesos de covariación cruzada de las martingalas formadas a partir de M .

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 $K(A \times B \times C) := \lambda^d(A \cap B) \lambda(C)$.

Ya hemos visto que la integral de Itô-Walsh es una medida martingala para funciones elementales y comentamos que lo mismo sucede para funciones simples. Más aún, resulta ser que si M es una medida martingala digna, entonces la integral de Itô-Walsh de funciones simples también es una medida martingala digna. Este es el contenido del siguiente resultado.

Teorema 6. Sea M una medida martingala digna y f una función simple. Entonces $(f \cdot M)$ es una medida martingala digna. Mas aún, si Q_M y K_M representan al funcional de covariación y a la medida dominante de la medida martingala digna M , entonces

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

La prueba la omitimos, pero se puede consultar en [10, proposición 5.23].

A partir de ahora, solo estaremos interesados en el caso cuando la variable temporal está en algún intervalo finito $(0, T]$.

Hasta ahora hemos definido la integral de Itô-Walsh en el conjunto \mathcal{S} de las funciones simples, y hemos probado propiedades básicas de la misma. Siguiendo la maquinaria estándar para definir nuevos objetos a partir de aproximaciones, el siguiente paso natural es extender la definición a todo un conjunto \mathcal{P} de funciones predecibles, para alguna medida martingala M . Dotamos primero a dicho conjunto de una norma apropiada para dicho propósito.

Definición 12. Sea M una medida martingala digna. Supongamos que K_M es su medida dominante. Sea $f \in \mathcal{P}$ una función predecible. Definimos la norma $\|\cdot\|_M$ en dicho espacio 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 pues K_M es una medida signada definida positiva. Denotamos por \mathcal{P}_M al conjunto de funciones predecibles f tales que $\|f\|_M < \infty$.

Teorema 7. La norma anteriormente definida en efecto es una norma, y el conjunto \mathcal{P}_M es completo con esta norma. Por lo tanto, el espacio $(\mathcal{P}_M, \|\cdot\|_M)$ es un espacio de Banach.

Demostración. Suponer que la norma es idénticamente cero es equivalente a decir que

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

lo cual sólo es posible cuando $|f(x, t)f(y, t)| = 0$, independientemente de x, y y t . Esto se cumple sólo cuando f es la función cero salvo conjuntos de medida K_M igual a cero. El producto por escalar claramente se factoriza fuera de la norma por las propiedades del valor absoluto y de las integrales.

Para la parte de la desigualdad del triángulo, dado que \mathcal{P}_M es por definición el conjunto de todas las funciones medibles con respecto a una σ -álgebra, utilizamos el lema de Dynkin funcional: primero mostramos la propiedad para las funciones elementales, luego probamos la propiedad en \mathcal{S} (las generadoras del espacio \mathcal{P}_M) y posteriormente extendemos con Dynkin. El argumento completo se omite en el texto.

Finalmente, para probar que el espacio $(\mathcal{P}_M, \|\cdot\|_M)$ es completo, usamos [18, proposición 2.2] y concluimos. ■

Ya que el espacio $(\mathcal{P}_M, \|\cdot\|_M)$ tiene una estructura adecuada, construimos un «puente» para extender la integral de \mathcal{S} a todo \mathcal{P}_M . Ese es el objetivo de la siguiente proposición.

Teorema 8. El conjunto \mathcal{S} es denso en \mathcal{P}_M .

Para una prueba de este hecho remitimos a [18, proposición 2.3]. Lo anterior junto con la proposición 5 y las propiedades de la medida dominante K_M nos dice que para rectángulos,

$$\mathbb{E} [(f \cdot M)_t^2(B)] \leq \|f\|_M^2, \quad \text{para } t \in (0, T], \quad f \in \mathcal{S}, \quad B \in \mathcal{B}(\mathbb{R}^d). \quad (1.8)$$

Armados con la cota anterior, procedemos a extender la integral a una función $f \in \mathcal{P}_M$ de la siguiente forma. Dado que $\mathcal{S} \subseteq \mathcal{P}_M$ es un conjunto denso, para $f \in \mathcal{P}_M$ existe una sucesión de funciones $(f_n)_{n \geq 1} \subseteq \mathcal{S}$ tales que $\|f_n - f\|_M \xrightarrow{n \rightarrow \infty} 0$. Esto hace a $(f_n)_{n \geq 0}$ una sucesión de Cauchy. Luego, gracias a la cota (1.8), la sucesión $((f_m \cdot M)_t(B))_{m \geq 1}$ también es de Cauchy en $L^2(\mathbb{P})$. Obsérvese que los elementos de esta última sucesión están bien definidos gracias a que la integral de Itô-Walsh está bien definida en \mathcal{S} . Dado que $L^2(\mathbb{P})$ es completo, tenemos la existencia de una martingala, que denotaremos por $(f \cdot M)_t(B)$, de tal forma que

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

y dicho objeto no depende de la sucesión $(f_n)_{n \geq 1}$ que aproxima a $f \in \mathcal{P}_M$. Dicho objeto es, por definición, la integral de Itô-Walsh para una función f predecible. Todo lo anterior nos lleva al siguiente resultado.

Teorema 9. Sea M una medida martingala digna. Entonces para cualquier $f \in \mathcal{P}_M$, el objeto $(f \cdot M)$ es una medida martingala digna que satisface las propiedades de la proposición 6. 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 $L^2(\mathbb{P})$.

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

La prueba del teorema anterior se puede consultar en [18, teorema 2.5]. A partir de aquí, adoptamos la siguiente notación para la integral de Itô-Walsh, la cual será usada principalmente en el capítulo 4:

$$\int_{A \times [0, t]} f(x, s) M(dx, ds) := (f \cdot M)_t(A), \quad \text{para } f \in \mathcal{P}_M, A \in \mathcal{B}(\mathbb{R}^d) \text{ y } t > 0.$$

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 10 (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 $t > 0$,

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

Para una prueba de un caso especial, se puede consultar [10, teorema 5.27].

A menudo será de utilidad intercambiar el orden de integración cuando una integral de Lebesgue «usual» y una integral de Itô-Walsh conviven. Este es el contenido del siguiente resultado.

Teorema 11 (Fubini estocástico). Sea (G, \mathcal{G}, μ) un espacio de medida finita y sea M una medida martingala con medida dominante K_M . Sea $f(x, s, \omega, \lambda)$ una función $\mathcal{P}_M \times \mathcal{G}$ medible, donde $x \in \mathbb{R}^d$, $s \geq 0$, $\omega \in \Omega$ y $\lambda \in G$. Supongamos que se cumple la condición

$$\mathbb{E} \left[\int_{\mathbb{R}^d \times \mathbb{R}^d \times [0, T] \times G} |f(x, s, \omega, \lambda) f(y, s, \omega, \lambda)| K_M(dx, dy, ds) \mu(d\lambda) \right] < \infty.$$

Entonces

$$\int_G \left[\int_{\mathbb{R}^d \times [0, t]} f(x, s, \lambda) M(dx, ds) \right] \mu(d\lambda) = \int_{\mathbb{R}^d \times [0, t]} \left[\int_G f(x, s, \lambda) \mu(d\lambda) \right] M(dx, ds).$$

La prueba de este resultado se puede hallar en [18, teorema 2.6].

Con esto terminamos la construcción de la integral de Itô-Walsh con respecto a una medida martingala. Para finalizar esta sección, presentamos como ejemplo a la integral de Itô-Walsh con respecto a un ruido blanco.

Ejemplo 6. Sea \dot{W} un ruido blanco en $\mathbb{R}^d \times [0, \infty)$. Sea $(W_t(A))_{t \geq 0, A \in \mathcal{B}(\mathbb{R}^d)}$ el proceso de ruido blanco asociado al mismo. Ya hemos visto que dicho proceso forma una medida martingala digna con respecto a la filtración (1.6).

Entonces de acuerdo al teorema 9, para una función f predecible con respecto a la filtración generada por el proceso de ruido blanco, existe una nueva medida martingala digna en $L^2(\mathbb{P})$, denotada

$$(f \cdot W)_t(A) = \int_{A \times [0, t]} f(x, s) W(dx, ds), \quad \text{para } A \in \mathcal{B}(\mathbb{R}^d), t \geq 0.$$

Dicho objeto es la integral de Itô-Walsh de f con respecto a un ruido blanco \dot{W} .

Ahora que ya nos es posible hablar de integrales con respecto a un ruido blanco \dot{W} , es posible estudiar de manera rigurosa lo que significa resolver una ecuación diferencial parcial estocástica cuya componente aleatoria está dada por un ruido blanco.

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

Como vimos en la introducción del capítulo, la heurística para reformular un problema de ecuaciones diferenciales parciales estocásticas de tal forma que se logre «evadir» el uso de las derivadas convencionales, pero que sea equivalente al problema original, 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.

Para ejemplificar lo anterior, presentamos un ejemplo de reformulación a una ecuación diferencial parcial estocástica conocida: la ecuación del calor estocástica en $[0, L] \times [0, \infty)$. Este ejemplo se puede consultar con más detalle en [10, pp. 23 - 32]. Consideremos el siguiente problema del calor,

$$\begin{cases} \partial_t u(x, t) = \partial_{xx} u(x, t) + f(u(x, t)) \dot{W}(x, t), & \text{si } t > 0, x \in [0, L] \\ \partial_x u(0, t) = \partial_x u(L, t) = 0, & \text{si } t > 0, \\ u(x, 0) = u_0(x), & \text{si } x \in [0, L], \end{cases} \quad (1.9)$$

donde \dot{W} es un ruido blanco con respecto a alguna filtración $(\mathcal{F}_t)_{t \geq 0}$ continua por derecha, 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 la ecuación diferencial 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 al ser u un objeto aleatorio no necesariamente será derivable. Por lo tanto, si de manera formal utilizamos integración por partes, al ser ϕ una función cuya derivada

se anula en las fronteras, la integral de la parcial de u multiplicado por ϕ se transforma en

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

Sustituyendo la expresión anterior en la reformulación 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, resolver dicha reformulación integral se puede considerar como resolver el problema original. Estas maneras de resolver una ecuación diferencial parcial estocástica varían, dependiendo de la naturaleza del problema a tratar. Las soluciones a estas ecuaciones halladas con la reformulación del problema en términos de ecuaciones integrales se conocen como *soluciones mild*.

En el ejemplo que presentamos, la solución mild al problema del calor (1.9) se establece con rigor en la siguiente definición.

Definición 13. Decimos que u es una solución *mild* del problema inicial anterior, si existe un proceso estocástico (también llamado *campo aleatorio*) $u(x, t)$ medible y adaptado a la filtración generada por W tal que para cualquier función $\phi \in C^\infty([0, L])$, donde $\phi'(0) = \phi'(L) = 0$, se tiene la igualdad

$$\begin{aligned} \int_0^L u(x, t) \phi(x) dx - \int_0^L u_0(x) \phi(x) dx &= \int_0^t \int_0^L u(x, s) \phi''(x) dx ds \\ &+ \int_0^t \int_0^L f(u(x, s)) \phi(x) W(dx, ds). \end{aligned}$$

Se conoce en la literatura la solución del problema (1.9). Presentamos aquí el teorema correspondiente.

Teorema 12. El problema del calor (1.9), con las condiciones especificadas anteriormente sobre \dot{W} , f y u_0 , tiene una única solución \mathbb{P} - casi seguramente, que satisface lo siguiente para cualquier $T > 0$.

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

Mas aún, el siguiente teorema nos muestra que no solamente hay existencia y unicidad de las solución al problema (1.9), sino que también existe una modificación continua de la misma.

Teorema 13. Sea $\tilde{u}(x, t)$ la solución única \mathbb{P} - casi seguramente del problema (1.9). Entonces existe una modificación $u(x, t)$ de dicha solución cuyas trayectorias son continuas.

La prueba de ambos resultados se puede encontrar en [10, teorema 6.4, teorema 6.7], respectivamente.

Para concluir el capítulo y en preparación para el problema central de la tesis, estamos interesados la solución al problema del calor estocástico siguiente.

$$\begin{cases} \partial_t u = \frac{1}{2} \partial_{xx} u + \sigma(u) \dot{W} & \forall x \in \mathbb{R}, \forall t > 0, \\ u_0(x) := u(x, 0) = 1 & \forall x \in \mathbb{R}, \end{cases} \quad (1.10)$$

con \dot{W} un ruido blanco definido en $[0, \infty) \times \mathbb{R}$, el cual podemos interpretar como un ruido blanco espacio-temporal, y σ una función globalmente Lipschitz tal que $\sigma(1) \neq 0$.

Aunque la solución al problema anterior será fuertemente usada, dado que no es el propósito principal del texto, presentamos sin demostración el resultado que nos garantiza la existencia y unicidad del problema de calor anterior. El resultado que presentaremos está tomado de [11, proposición 2.1], pero una versión general junto con su prueba se pueden encontrar en [3, teorema 2.4]. Para enunciar el teorema que nos interesa, debemos otorgar algo de contexto, el cual también es extraído de [3].

Consideremos primero la solución fundamental al problema de calor clásico en $\mathbb{R} \times [0, \infty)$,

$$\begin{cases} \partial_t u = \frac{\nu}{2} \partial_{xx} u, & \forall x \in \mathbb{R}, \forall t > 0, \nu > 0. \\ u(x, 0) = \delta(x), & \forall x \in \mathbb{R}. \end{cases}$$

donde $\delta(x)$ es la delta de Dirac. En nuestro caso, estamos interesados simplemente en $\nu = 1$. Dicha solución fundamental se conoce que está dada por el núcleo de calor, que denotamos por

$$p_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}, \quad x \in \mathbb{R}, t \geq 0.$$

Dentemos por $J_0(t, x)$ la solución al problema de calor pero ahora con condición inicial μ y $\nu = 1$,

$$\begin{cases} \partial_t u = \frac{1}{2} \partial_{xx} u, & \forall x \in \mathbb{R}, \forall t > 0, \\ u(x, 0) = \mu(x), & \forall x \in \mathbb{R}. \end{cases}$$

donde μ es una medida signada de Borel sobre todo \mathbb{R} tal que

$$\int_{\mathbb{R}} e^{-ax^2} |\mu|(dx) < \infty, \quad \forall a > 0. \quad (1.11)$$

Dicha solución fundamental se conoce que está dada por la convolución de μ con el núcleo del calor, a saber,

$$J_0(t, x) = (\mu * p_t(\cdot))(x) = \int_{\mathbb{R}} p_t(x - y) \mu(dy).$$

Entonces la ecuación del calor del problema estocástico (1.10) con condición inicial $u(x, 0) = \mu(x)$ (no necesariamente $u(0, x) \equiv 1$) se puede reescribir combinando la solución a la ecuación

homogénea J_0 y la integral del término estocástico, obteniendo al menos de manera formal, la siguiente forma integral,

$$\begin{aligned} u(t, x) &= J_0(t, x) + I(t, x) \\ &= \int_{\mathbb{R}} p_t(x - y) \mu(dy) + \int_{[0, t] \times \mathbb{R}} p_{t-s}(x - y) \sigma(u) W(ds, dy). \end{aligned} \quad (1.12)$$

Por lo tanto, una solución mild a la ecuación del calor estocástica sobre $\mathbb{R} \times [0, \infty)$ pero con condición inicial $u(0, x) = \mu(x)$ una medida signada de Borel en \mathbb{R} que cumple la condición (1.11), debería cumplir la ecuación (1.12). En particular, si consideramos $\mu(x) \equiv 1$, directamente tenemos que $J_0(t, x) = 1$, como es el caso que nos interesa. Llegamos así al siguiente teorema, el cual se enuncia en concordancia con [11, proposición 2.1].

Teorema 14. Consideremos el problema de calor (1.10). Entonces existe un único campo aleatorio $u = \{u(t, x) : (t, x) \in [0, \infty) \times \mathbb{R}\}$ medible y adaptado, tal que para cualquier $T > 0$ y $p \geq 2$,

$$\sup_{(t, x) \in [0, T] \times \mathbb{R}} \mathbb{E}[|u(t, x)|^p] = C_{T, p}, \quad (1.13)$$

y para cualquier $t \geq 0$ y $x \in \mathbb{R}$, u satisface

$$u(t, x) = 1 + \int_{[0, t] \times \mathbb{R}} p_{t-s}(x - y) \sigma(u(s, y)) W(ds, dy). \quad (1.14)$$

Concluimos aquí nuestro breve estudio de ecuaciones diferenciales parciales estocásticas. Retomaremos estos conceptos en el capítulo 4, cuando la teoría necesaria para abordar el problema principal de la tesis esté cubierta.

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. Más allá de 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 [14]), 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, forman 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 del proceso de Ornstein-Uhlenbeck. 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 [14], 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 llevar 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 que resultan de gran utilidad. Tal es el caso del siguiente lema.

Lema 1. (Identidad de Stein) 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,

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

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 cálculos 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 $f(0) \neq 0$, utilizamos la función $g(x) = f(x) - f(0)$ y repetimos las cuentas. ■

En la búsqueda de definir la derivada para funciones más generales, definimos a la clase de Schwarz \mathcal{S} como el conjunto de funciones en $C^\infty(\mathbb{R}, \mathbb{R})$ con derivadas de crecimiento a lo más polinomial, es decir,

$$\mathcal{S} := \left\{ f \in C^\infty(\mathbb{R}) : \forall p \in \{0, 1, 2, \dots\}, \forall n \geq 1, \text{ se satisface } \sup_{x \in \mathbb{R}} (1 + |x|^n) |f^{(p)}(x)| < \infty \right\}.$$

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 supremos. 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 (ver, por ejemplo, [1, corolario 2.3.25]). 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. ■

Sabiendo que \mathcal{S} es un subespacio lineal denso de $L^q(\gamma)$, tenemos un marco teórico por medio del cual generalizar la derivada de las funciones en el espacio de Schwarz a funciones cualesquiera de $L^q(\gamma)$. 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. Buscando 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} . Esta idea se formaliza con el concepto de *operador cerrable* (ver, por ejemplo, [1, pp. 62]). En nuestro caso, la derivada de orden p en efecto es un operador cerrable. Este es el contenido del siguiente teorema.

Teorema 15. 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.

Probamos solamente 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 $D^p f = \frac{dp}{dx^p} f$ para cualquier $f \in \mathcal{S}$. Estos elementos nos permiten definir la derivada de Malliavin.

Definición 14. Sea $p \geq 1$ un entero positivo. Definimos la *derivada de Malliavin* como el operador D^p que extiende a la derivada $\frac{dp}{dx^p}$ vista como operador de \mathcal{S} a $L^q(\gamma)$, para $q \in [1, \infty)$. Referenciar a la norma definida abajo. En el peor de los casos repito el outcome del teorema de funcional

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.

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 $\|\cdot\|_{\mathbb{D}^{p,q}} : \mathbb{D}^{p,q} \rightarrow [0, \infty)$ 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 que 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.

El espacio $\mathbb{D}^{p,q}$ cumple las siguientes relaciones de contención.

Teorema 16. Sean $p \geq 0$ y $q' > q$. Entonces se cumple que

$$\mathbb{D}^{p+1,q'} \subseteq \mathbb{D}^{p,q},$$

además de que por convención, $\mathbb{D}^{0,q} = L^q(\Omega)$ y $\|\cdot\|_{0,q} = \|\cdot\|_q$, y denotaremos

$$\mathbb{D}^{\infty,q} := \bigcap_{p \geq 0} \mathbb{D}^{p,q}.$$

Más aún, la derivada D^p es compatible en la intersección $\mathbb{D}^{p,q} \cap \mathbb{D}^{p,q'}$, es decir, para una función f en la intersección $\mathbb{D}^{p,q} \cap \mathbb{D}^{p,q'}$ la derivada de orden p como operador con dominio $\mathbb{D}^{p,q}$ aplicado a f , es el mismo objeto que la derivada de orden p como operador con dominio $\mathbb{D}^{p,q'}$ aplicado a f . [Poner en palabras](#)

La prueba de este teorema se puede encontrar en [15, p. 27]. En particular el resultado nos dice que, al ser $\mathbb{D}^{p,q}$ subconjuntos de los espacios $L^q(\Omega)$, y ser este último un espacio de medida finita, entonces la contención $L^{q'}$ para $q' > q$ se sigue preservando en los espacios de Watanabe-Sobolev. La contención en la variable p nos dice simplemente que si una variable es derivable $n+1$ veces, claramente será n veces derivable.

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. En preparación para esto, tenemos el siguiente resultado

Teorema 17. Sea $p \geq 1$ y consideremos el espacio $\mathbb{D}^{p,2}$. Dicho espacio es un espacio de Hilbert con producto interno dado por

$$\langle f, g \rangle_{\mathbb{D}^{p,2}} := \mathbb{E}[fg] + \sum_{k=1}^p \mathbb{E}[\langle D^k f, D^k g \rangle_{\mathfrak{H}^{\otimes k}}].$$

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, 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 $c_g > 0$ que cumple

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

entonces para cada $g \in \text{Dom}(\delta^p)$ el funcional lineal, con dominio $\mathbb{D}^{2,p}$ y con regla de correspondencia $f \longmapsto \langle D^p f, g \rangle$, 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 15. Sea $p \geq 1$ un entero positivo. Denotamos

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

y definimos el operador de divergencia de orden p como la aplicación $\delta^p : \text{Dom}(\delta^p) \longrightarrow L^2(\gamma)$ 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 del operador no acotado 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,

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

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 como adjunto de D , se tiene que [escribir que utilicé la propiedad de adjunto en g y la mandé del lado izquierdo](#)

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

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

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

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

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

2. 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,

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

3. 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

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

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.

Ya que \mathcal{S} es un subconjunto denso de $L^q(\gamma)$ y este último espacio es completo, 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 19. 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 bajo γ , el espacio tiene medida finita, basta con probar el resultado para $L^1(\gamma)$. Sea $f \in L^1(\gamma)$. Tenemos las siguientes igualdades para $x \in \mathbb{R}$:

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

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

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

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

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

■

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 20. Sea $f \in \mathbb{D}^{1,2}$ y $t \geq 0$. Entonces $P_tf \in \mathbb{D}^{1,2}$ y $DP_tf = e^{-t}P_tDf$.

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

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

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 la familia de operadores $(P_t)_{t \geq 0}$ con el semigrupo de Ornstein-Uhlenbeck:

Teorema 21. 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 $(P_t)_{t \geq 0}$.

Al proceso X_t^x que resuelve la ecuación anterior se le conoce como *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:

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

■

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

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,

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

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 conjunto 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

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

Cuando estemos hablando del generador infinitesimal aplicado a funciones pertenecientes a un conjunto particular, omitiremos escribir la función f en el generador. 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 22. 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

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

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,

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

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 el lema 3, 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 23. (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 la versión de Cálculo de Malliavin de la desigualdad.

Teorema 24. (Desigualdad de Poincaré unidimensional) Sea $N \sim \text{Normal}(0, 1)$ y sea $f \in \mathbb{D}^{1,2}$. Entonces

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

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:

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

de manera que, si $\mathbb{E}[f(N)] = 0$, directamente recuperamos la desigualdad de Poincaré clásica, aunque en distinto contexto. 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 25. 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 requisitos de [11, sección 2.1]. Dado que el caso unidimensional fue tratado con relativo detalle, y nuevamente teniendo en cuenta que un estudio exhaustivo de la disciplina no es el objetivo de este texto, las pruebas de los resultados que aparecen en el resto del capítulo serán omitidas, por lo que remitimos a [14, capítulo 2] para su consulta.

Para poder hablar del cálculo de Malliavin en dimensiones infinitas, necesitamos un contexto adecuado. Parte de ese contexto es el concepto de *proceso gaussiano isonormal*. En el ejemplo 3 del capítulo 1 hablamos del proceso gaussiano isonormal inducido por un ruido blanco. A continuación presentamos la definición de un proceso gaussiano isonormal en un contexto general.

Definición 17. Sean $(\mathfrak{H}, \langle \cdot, \cdot \rangle_{\mathfrak{H}})$ un espacio de Hilbert separable sobre \mathbb{R} , y $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Un *proceso gaussiano isonormal* sobre \mathfrak{H} es un proceso gaussiano centrado $X = \{X(h) : h \in \mathfrak{H}\}$ definido en Ω , tal que para cualesquiera $g, h \in \mathfrak{H}$,

$$\mathbb{E}[X(g)X(h)] = \langle g, h \rangle_{\mathfrak{H}}.$$

Dado un proceso gaussiano isonormal X , al ser sus elementos variables conjuntamente gaussianas y centradas, el proceso está definido completamente por su función de covarianzas, y esta se codifica por medio del producto interno del espacio de Hilbert \mathfrak{H} . Es por lo tanto una manera de relacionar íntimamente dos estructuras distintas. De manera más precisa, consideremos un sistema ortonormal $(Z_n)_{n \geq 1}$ en el espacio de Hilbert \mathfrak{H} . Sea $G = \{X(Z_n) : n \geq 1\}$ una familia de variables gaussianas independientes centradas en un espacio de probabilidad

adecuado. Consideremos a $L^2(G)$ el subespacio lineal cerrado del $L^2(\mathbb{P})$ generado por dichas variables. Entonces el mapeo $h \rightarrow X(h)$ resulta ser una isometría lineal del espacio H en $L^2(G)$. En particular, dos procesos gaussianos isonormales definidos sobre el mismo espacio de Hilbert tendrán la misma ley.

Las ideas anteriormente expuestas, en conjunto con el teorema de consistencia de Kolmogorov, son esencialmente la prueba de la existencia de un proceso isonormal dado un espacio de Hilbert \mathfrak{H} . Esto se establece en el siguiente resultado.

Teorema 26. Sea $(\mathfrak{H}, \langle \cdot, \cdot \rangle_{\mathfrak{H}})$ un espacio de Hilbert real y separable. Entonces existe un proceso gaussiano isonormal sobre \mathfrak{H} .

Dado que cualquier espacio de Hilbert puede ser la base para construir un proceso gaussiano isonormal, en preparación para el capítulo 4, podemos pensar en $\mathfrak{H} = L^2([0, \infty) \times \mathbb{R})$, esto es, el espacio de todas las funciones de dos variables $f(t, x)$ que son Lebesgue cuadrado-integrables. Este espacio será el que usaremos en el capítulo cuatro para el resultado principal. Ahora denotaremos por \mathcal{S} al siguiente conjunto de funciones:

$$\mathcal{S} := \{f(X(h_1), \dots, X(h_n)) : f \in C_p^\infty(\mathbb{R}^n, \mathbb{R}), h_1, \dots, h_n \in \mathfrak{H}\},$$

donde $X = \{X(h) : h \in \mathfrak{H}\}$ es el proceso gaussiano isonormal sobre el cual estamos trabajando, y $C_p^\infty(\mathbb{R}^n, \mathbb{R})$ es el conjunto de funciones infinitamente derivables tales que tanto ella como sus derivadas parciales tienen crecimiento a lo más polinomial. Dicha clase \mathcal{S} de funciones se conocen en general como *funciones suaves*, y son el conjunto análogo a la clase de Schwarz que se manejó en la sección anterior en el caso unidimensional.

Presentamos a continuación una familia de espacios particular.

Definición 18. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $(\mathfrak{H}, \langle \cdot, \cdot \rangle_{\mathfrak{H}})$ un espacio de Hilbert real y separable. Definimos

$$L^q(\Omega; \mathfrak{H}) := \{Y(\omega) \in \mathfrak{H} : Y \text{ es } \mathcal{F}\text{-medible, } \mathbb{E} [\|Y\|_{\mathfrak{H}}^q] < \infty, \omega \in \Omega\}.$$

En otras palabras, $L^q(\Omega; \mathfrak{H})$ representa el conjunto de *elementos aleatorios* Y que toman valores en \mathfrak{H} , y que vistos como funciones de (Ω, \mathcal{F}) en $(\mathfrak{H}, \mathcal{B}(\mathfrak{H}))$, son medibles, y cuya norma en \mathfrak{H} tiene momentos de orden q finitos.

Llegados a este punto, estamos listos para definir la derivada de Malliavin en el caso general.

Definición 19. Sea $F \in \mathcal{S}$ una función suave, y sea $p \geq 1$ un entero. La derivada de Malliavin de orden p de F (con respecto al proceso isonormal X) es el elemento de $L^2(\Omega; \mathfrak{H}^{\odot p})$ definido por

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

en donde $\mathfrak{H}^{\odot p}$ representa el *producto tensorial simetrizado* de orden p del espacio de Hilbert \mathfrak{H} consigo mismo (el cual nuevamente es un espacio de Hilbert). En el caso en que $p = 1$ hablaremos simplemente de la *derivada de Malliavin* de F y la denotamos por DF .

Observación 2. Nótese que para $F \in \mathcal{S}$ y $h \in \mathfrak{H}$ arbitrarios, al ser $DF \in L^2(\Omega; \mathfrak{H})$ una variable aleatoria con valores en \mathfrak{H} ,

$$\langle DF, h \rangle_{\mathfrak{H}} = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (f(X(h_1) + \varepsilon \langle h_1, h \rangle, \dots, X(h_n) + \varepsilon \langle h_n, h \rangle) - F),$$

por lo que el producto interno entre DF y h se puede interpretar como *una derivada direccional*. Por tal motivo, denotaremos tal producto interior como

$$D_h F := \langle DF, h \rangle_{\mathfrak{H}}.$$

La derivada de Malliavin, vista como operador entre espacios, también cumple la propiedad de ser cerrable, tal y como era de esperarse en concordancia con el caso unidimensional.

Teorema 27. 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, podemos replicar los pasos hechos para el caso unidimensional y obtener un espacio adecuado en donde dicho operador sea cerrado: para $q \in [1, \infty)$ y $p \geq 1$ un entero, definimos la norma $\|\cdot\|_{\mathbb{D}^{p,q}} : \mathcal{S} \rightarrow [0, \infty)$,

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

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

Como en el caso unidimensional, en el caso general se tiene un teorema análogo al teorema 16. Omitimos su enunciado y su prueba. También se tiene que para cualquier $p \geq 1$, el espacio $\mathbb{D}^{p,2}$ es un espacio de Hilbert con producto interno análogo al del caso unidimensional.

Un resultado importante y que será usado posteriormente es la *regla de la cadena* para la derivada de Malliavin. Éste se encuentra en el siguiente teorema.

Teorema 28 (Regla de la cadena). Sea $\phi : \mathbb{R}^m \rightarrow \mathbb{R}$ una función C^1 con derivadas parciales acotadas. Supongamos que $F = (F_1, \dots, F_n)$ es un vector aleatorio cuyas componentes son elementos de $\mathbb{D}^{1,q}$ para algún $q \geq 1$. Entonces $\phi(F) \in \mathbb{D}^{1,q}$ y

$$D\phi(F) = \sum_{i=1}^n \frac{\partial \phi}{\partial x_i}(F) DF_i.$$

Pasamos ahora a estudiar el operador de divergencia. Fijamos $q = 2$. Dado un entero $p \geq 1$, definimos el operador de divergencia de orden p como el operador adjunto del operador derivada $D^p : (\mathbb{D}^{p,2}, \|\cdot\|_{\mathbb{D}^{p,q}}) \rightarrow (L^2(\Omega, \mathfrak{H}^{\otimes p}), \|\cdot\|_{2, \mathfrak{H}^{\otimes p}})$.

Definición 20. Sea $p \geq 1$ un entero. Denotamos por $\text{Dom}(\delta^p)$ al subconjunto de $L^2(\Omega; \mathfrak{H}^{\odot p})$ 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 21. 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*. Tal y como sucede en el caso unidimensional, al ser el operador de divergencia el adjunto del operador derivada D^p y ser este último cerrado, se tiene que δ^p es un operador cerrado.

El siguiente resultado nos permite «sacar» una variable aleatoria escalar dentro del operador de divergencia δ . Este resultado se puede consultar en [14, proposición 2.5.4], sin embargo, enunciamos una versión general que se encuentra en [2, lema 1].

Teorema 29. 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 completamente análoga al caso unidimensional.

Teorema 30. Se cumple que

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

Múltiples propiedades con respecto al operador derivada y el operador de divergencia se pueden obtener, generalizando las propiedades del caso unidimensional. En particular, las versiones adecuadas del semigrupo de Ornstein-Uhlenbeck, y de su generador infinitesimal asociado, pueden ser enunciadas. No obstante, dado que para nuestros propósitos la teoría desarrollada hasta aquí es suficiente, remitimos [14, capítulo 2] para continuar el estudio de estos objetos.

Cabe destacar que estaremos interesados en el caso en que el proceso gaussiano isonormal subyacente es aquél inducido por un proceso de ruido blanco gaussiano $(W_t(A))_{t \in [0, \infty)}$, $A \in \mathcal{B}(\mathbb{R})$, el cual se puede construir a partir del espacio de Hilbert $\mathfrak{H} = L^2([0, \infty) \times \mathbb{R})$. Este será nuestro contexto para el capítulo 4. Concluimos nuestro breve estudio de la disciplina con una propiedad muy interesante e importante del operador de divergencia: precisamente cuando estamos en el caso en que $\mathfrak{H} = L^2([0, \infty) \times \mathbb{R})$, y el proceso isonormal subyacente es un proceso de ruido blanco gaussiano, el operador δ *coincide* con la integral de Itô-Walsh en $[0, \infty) \times \mathbb{R}$.

Teorema 31. Sea u un proceso adaptado y cuadrado-integrable. Entonces $u \in \text{Dom}(\delta)$ y se cumple que

$$\delta(u) = \int_{[0, \infty) \times \mathbb{R}} u(s, y) W(ds, dy).$$

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. La exposición de este capítulo está basado principalmente en [14, capítulo 4].

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, está la propiedad de ser una ley determinada por sus momentos. Esta propiedad, la cual está codificada en el conocido como *lema de Stein*, forma parte importante de este texto.

Recordemos que, dada una medida ν en \mathbb{R} , el momento de orden n 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

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

en donde ξ y η son los puntos intermedios del residuo de Lagrange. Usando monotonía de la integral, la última expresión anterior es igual a

$$\begin{aligned} \int_{\mathbb{R}} \frac{|tx|^{n+1}}{(n+1)!} d\nu(x) + \int_{\mathbb{R}} \frac{|tx|^{n+1}}{(n+1)!} d\gamma(x) &\leq \left(\int_{\mathbb{R}} \frac{|tx|^{2n+2}}{(n+1)!^2} d\nu(x) \right)^{1/2} + \left(\int_{\mathbb{R}} \frac{|tx|^{2n+2}}{(n+1)!^2} d\gamma(x) \right)^{1/2} \\ &= \left(\frac{|t|^{2n+2}}{(n+1)!^2} m_{2n+2}(\nu) \right)^{1/2} + \left(\frac{|t|^{2n+2}}{(n+1)!^2} m_{2n+2}(\gamma) \right)^{1/2} \\ &= 2 \left(\frac{|t|^{2n+2}}{(n+1)!^2} m_{2n+2}(\gamma) \right)^{1/2}. \end{aligned}$$

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 32. (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 gaussiana. Por otro lado, un instante de reflexión nos puede llevar a la siguiente pregunta.

Supongamos que ahora X es una variable aleatoria y 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 $h \in L^1(\gamma)$.

Comenzamos definiendo una noción de distancia entre variables aleatorias. Recordamos el concepto de *clase separante de funciones*.

Definición 22. 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 $h \in \mathcal{H}$ se tiene 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 .

A continuación, hacemos rigurosa la idea de distancia entre las leyes de dos variables aleatorias con la siguiente definición.

Definición 23. 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

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

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

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 . De manera rigurosa, tenemos la siguiente definición.

Definición 24. 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 es igual Lebesgue casi en todas partes 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 33. 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

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

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 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 $e^{-x^2/2} f(x) \xrightarrow{|x| \rightarrow \infty} 0$ 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 33. Tomando esperanza a ambos lados de la ecuación de Stein, tenemos que f_h cumple que

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

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

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

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

$$\mathcal{H} = \{h : \mathbb{R} \rightarrow \mathbb{R} : h(x) = \mathbf{1}_{(-\infty, z]}(x), \quad z \in \mathbb{R}\}.$$

Teorema 34. Sea $z \in \mathbb{R}$ y consideremos a la función $h = \mathbf{1}_{(-\infty, z]}$ 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

$$d_{\text{Kol}}(X, N) \leq \sup_{f \in \mathcal{F}_{\text{Kol}}} |\mathbb{E}[f'(X)] - \mathbb{E}[Xf(X)]|, \quad (3.3)$$

en donde

$$\mathcal{F}_{\text{Kol}} := \left\{ f \in \text{AC}(\mathbb{R}) : f \text{ es derivable salvo un conjunto finito, } \|f\|_\infty \leq \frac{\sqrt{2\pi}}{4}, \|f'\|_\infty \leq 1 \right\},$$

donde $\text{AC}(\mathbb{R})$ es el conjunto de funciones absolutamente continuas sobre \mathbb{R} , y donde el supremo en (3.3) 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 [14]. Concluimos esta sección con el siguiente resultado, el cual nos da cotas explícitas para una familia de funciones específicas.

Teorema 35. Consideremos la familia de funciones

$$\mathcal{H} = \{h : \mathbb{R} \rightarrow \mathbb{R} : h(z) = \mathbf{1}_{\{z > x\}}z, \ x \in \mathbb{R}\}.$$

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 [8].

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 de variables independientes e idénticamente distribuidas con

media cero, varianza uno 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 34. Para el enunciado y demostración del mismo, remitimos a [14, sección 3.7].

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 la densidad 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 29 y ambos, por orden de aparición, se pueden encontrar en [2, proposición 1] y en [11, teorema 3.2].

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

Teorema 36. 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

$$f_F(x) = \mathbb{E} \left[\mathbb{1}_{\{F > x\}} \delta \left(\frac{u}{D_u F} \right) \right]. \quad (3.4)$$

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 [2, lema 3, lema 4].

Estamos interesados en aplicar tanto el lema anterior como el teorema 36 en condiciones que aparecerán en el capítulo 4. Por lo tanto, y siguiendo de cerca a [11, sección 3], tenemos el siguiente resultado.

Teorema 37. 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

$$\sup_{x \in \mathbb{R}} |f_F(x) - \Phi(x)| \leq (\|F\|_4 \| (D_v F)^{-1} \|_4 + 2) \|1 - D_v F\|_2 + \| (D_v F)^{-1} \|_4^2 \|D_v (D_v F)\|_2, \quad (3.5)$$

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 36 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 36, tenemos garantizada la existencia de una densidad continua y acotada para la variable $F = \delta(v)$, la cual está dada por la fórmula (3.4). Ahora buscamos una cota uniforme sobre $x \in \mathbb{R}$ para

$$|f_F(x) - \phi(x)| = \left| \mathbb{E} \left[\mathbf{1}_{\{F > x\}} \delta \left(\frac{u}{D_u F} \right) \right] - \mathbb{E} [\mathbf{1}_{\{N > x\}} N] \right|, \quad (3.6)$$

donde $N \sim \text{Normal}(0, 1)$. En virtud de que se satisfacen las hipótesis del lema 36, en la fórmula (3.4) podemos usar la ecuación (2.1) y obtenemos que, para $x \in \mathbb{R}$ arbitrario y omitiendo el subíndice \mathfrak{H} de los productos interiores,

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

donde hemos hecho uso de que $F = \delta(u)$ y de la regla de la cadena en la segunda igualdad. Sustituyendo esta última igualdad en la fórmula (3.6), y subsecuentemente sumando y restando el término $\mathbb{E} [\mathbf{1}_{\{F > x\}} F]$, obtenemos

$$\begin{aligned} |f_F(x) - \Phi(x)| &= \left| \mathbb{E} \left[\mathbf{1}_{\{F > x\}} \frac{F}{D_u F} \right] + \mathbb{E} \left[\mathbf{1}_{\{F > x\}} \frac{D_u D_u F}{(D_u F)^2} \right] - \mathbb{E} [\mathbf{1}_{\{N > x\}} N] \right| \\ &= \left| \mathbb{E} \left[\mathbf{1}_{\{F > x\}} \frac{F}{D_u F} \right] - \mathbb{E} [\mathbf{1}_{\{F > x\}} F] + \mathbb{E} \left[\mathbf{1}_{\{F > x\}} \frac{D_u D_u F}{(D_u F)^2} \right] \right. \\ &\quad \left. + \mathbb{E} [\mathbf{1}_{\{F > x\}} F] - \mathbb{E} [\mathbf{1}_{\{N > x\}} N] \right| \\ &\leq \left| \mathbb{E} \left[\mathbf{1}_{\{F > x\}} \left(\frac{F}{D_u F} - F \frac{D_u F}{D_u F} \right) \right] \right| + \left| \mathbb{E} \left[\mathbf{1}_{\{F > x\}} \frac{D_u D_u F}{(D_u F)^2} \right] \right| \\ &\quad + |\mathbb{E} [\mathbf{1}_{\{F > x\}} F] - \mathbb{E} [\mathbf{1}_{\{N > x\}} N]| \\ &\leq \mathbb{E} \left[\left| \frac{F(1 - D_u F)}{D_u F} \right| \right] + \mathbb{E} \left[\frac{|D_u D_u F|}{(D_u F)^2} \right] + |\mathbb{E} [\mathbf{1}_{\{F > x\}} F] - \mathbb{E} [\mathbf{1}_{\{N > x\}} N]|. \quad (3.7) \end{aligned}$$

Ahora, en la desigualdad (3.7) usamos sucesivamente la desigualdad de Hölder en el primer y segundo términos, de manera que

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

y también

$$\mathbb{E} \left[\frac{|D_u D_u F|}{(D_u F)^2} \right] \leq \|D_u D_u F\|_2 \|(D_u F)^{-2}\|_2 = \|D_u D_u F\|_2 \|(D_u F)^{-1}\|_4^2. \quad (3.9)$$

Finalmente, estudiamos el tercer término de la desigualdad (3.7). Aquí es donde utilizamos el método de Stein. Observemos que, para $x \in \mathbb{R}$,

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

donde $h(z) = \mathbb{1}_{\{z>x\}} z$ es una función perteneciente a la familia \mathcal{H} correspondiente al teorema 35. En virtud de la ecuación (3.2), la igualdad anterior se convierte en

$$|\mathbb{E} [\mathbb{1}_{\{F>x\}} F] - \mathbb{E} [\mathbb{1}_{\{N>x\}} N]| = |\mathbb{E} [f'_h(F)] - \mathbb{E} [F f_h(F)]|. \quad (3.10)$$

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 para la derivada de Malliavin, 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 (3.10) se sigue que

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

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

$$|\mathbb{E} [f'_h(F)(1 - D_v F)]| \leq \|f'_h(F)\|_2 \|1 - D_v F\|_2. \quad (3.11)$$

Finalmente, usamos el teorema 35, en (3.11), por lo que

$$|\mathbb{E} [\mathbb{1}_{\{F>x\}} F] - \mathbb{E} [\mathbb{1}_{\{N>x\}} N]| \leq \|C(|F| + 1)\|_2 \|1 - D_v F\|_2 \quad (3.12)$$

Finalmente, tenemos que $\|F\|_2^2 = \mathbb{E} [F^2] = 1$ y $C \leq 1$, por lo que (3.12) se convierte en

$$|\mathbb{E} [\mathbb{1}_{\{F>x\}} F] - \mathbb{E} [\mathbb{1}_{\{N>x\}} N]| \leq 2 \|1 - D_v F\|_2. \quad (3.13)$$

Combinando las desigualdades (3.8), (3.9) y (3.12), tenemos para $x \in \mathbb{R}$,

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

Tomando supremo obtenemos la desigualdad (3.5). ■

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 al problema del calor (1.10), a la densidad normal.

Hasta aquí llega nuestro estudio del método de Stein y de la existencia de la densidad de una variable aleatoria bajo ciertas condiciones de la misma. Es de destacar que el estudio realizado es solamente una pequeñísima parte de la vanguardia de ambos temas. Por un lado

las ecuaciones de Stein unidimensionales han sido ampliamente estudiadas, mientras que el estudio de existencia de densidades también se encuentra profundizado hoy en día, de acuerdo con [14].

En lo que a este texto respecta, la herramienta necesaria para abordar el problema planteado al inicio ha sido cubierta en su totalidad, por lo que en lo que resta, nos enfocaremos en estudiar dicho problema.

Capítulo 4

Propiedades espaciales de la Ecuación estocástica del calor

En este capítulo presentamos en su plenitud el problema planteado al inicio de este texto. Ahora que contamos con las herramientas necesarias para abordar nuestro problema, procedemos a enunciarlo de manera rigurosa. Este capítulo está completamente basado en [11, pp. 68-71, 75-86]. Consideremos nuevamente el problema del calor (1.10),

$$\begin{cases} \partial_t u = \frac{1}{2} \partial_{xx} u + \sigma(u) \dot{W}, & \forall x \in \mathbb{R}, \forall t > 0, \\ u_0(x) := u(x, 0) = 1, & \forall x \in \mathbb{R}, \end{cases}$$

donde \dot{W} es un ruido blanco en $[0, \infty) \times \mathbb{R}$, y σ es una función Lipschitz tal que $\sigma(1) \neq 0$.

Nos referiremos a este como nuestro problema principal, y siempre que hagamos referencia a él, entenderemos que las hipótesis sobre u_0 y σ anteriormente mencionadas se satisfacen. También, a lo largo de este capítulo, escribiremos a la densidad de una variable gaussiana centrada y de varianza $t > 0$ como

$$p_t(x) := \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}, \quad \forall x \in \mathbb{R}.$$

Del capítulo 1 tenemos garantizada la existencia y unicidad de un campo aleatorio $u = \{u(t, x) : (t, x) \in [0, \infty) \times \mathbb{R}\}$ que funge como solución al problema anterior.

Como mencionamos en dicho capítulo, diremos que u es la *solución mild* al problema (1.10). Con dicha solución al alcance, la cual es de naturaleza estocástica, nos interesamos ahora por el comportamiento asintótico de un nuevo objeto construido a partir de la misma: los *promedios espaciales*. Estos objetos, en nuestro contexto, se definen de la siguiente manera:

Definición 25. Sea $u = \{u(t, x) : (t, x) \in [0, \infty) \times \mathbb{R}\}$ la solución mild al problema (1.10). Para $R > 0$ fijo, definimos los promedios espaciales centrados y normalizados de u como

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

Nos referiremos a las variables $F_{R,t}$ simplemente como los promedios espaciales de la solución u , entendiéndolos siempre en el sentido de la definición anterior.

Tal y como se mencionó al inicio, en [9], para un $t > 0$ fijo, Huang, Nualart y Viitasaari estudian el comportamiento límite de las variables $F_{R,t}$ conforme R tiende a infinito, logrando probar resultados tipo teorema central del límite. Más aún, en dicha publicación, utilizando método de Stein y cálculo de Malliavin, 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, los autores obtuvieron el siguiente resultado.

Teorema 38. Sean u la solución mild al problema (1.10) y $Z \sim \text{Normal}(0,1)$. Supongamos que $\sigma_{R,t} > 0$. Entonces existe una constante $C > 0$ que depende solamente de t tal que

$$d_{\text{TV}}(F_{R,t}, N) \leq \frac{C}{\sqrt{R}}, \quad \text{para } t > 0 \text{ y } R \geq 1, \quad (4.1)$$

para alguna $C > 0$, independiente de R .

En particular, los autores de [9] resaltan que la condición $\sigma(1) \neq 0$ que aparece en el problema (1.10) implica que $\sigma_{R,t} > 0$.

Con este trabajo como antecedente, el propósito final del capítulo es obtener cotas superiores pero ahora 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, el resultado principal que se busca probar en este capítulo es el siguiente.

Teorema 39. Sea $u = \{u(t, x) : (t, x) \in [0, \infty) \times \mathbb{R}\}$ la solución mild al problema (1.10). Supongamos que $\sigma : \mathbb{R} \rightarrow \mathbb{R}$, además de las hipótesis propias del problema del calor, cumple las siguientes tres condiciones.

- $\sigma \in C^2(\mathbb{R})$.
- 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 tiene que para cualquier $R > 0$, los promedios espaciales $F_{R,t}$ poseen una densidad continua y acotada $f_{F_{R,t}}$, y la distancia uniforme entre las densidades de los promedios espaciales y la densidad de la distribución normal satisface la siguiente cota.

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

donde Φ es la densidad de una variable aleatoria normal estándar y $C_t > 0$ es una constante dependiente de t y de σ , pero independiente de R .

Este capítulo, dedicado a la prueba de este último teorema, consiste en dos secciones. La primera la dedicamos a mostrar dos resultados clave para la prueba. El primero de ellos tiene que ver con una expresión para la segunda derivada de Malliavin de la solución al problema del calor (1.10), mientras que el segundo tiene que ver con la existencia de los momentos inversos de la derivada de la misma. La segunda sección consiste en el cuerpo de la prueba como tal, y al inicio de ésta se describe como está estructurada.

4.1. Cotas para la primera y segunda derivada de la solución mild y existencia de momentos inversos

Estudiamos primero el comportamiento de la segunda derivada de la solución mild u a la ecuación del calor (1.10). Cabe mencionar que para lograrlo debemos basarnos en un resultado análogo para la primera derivada de u . Esto debido a que a partir del resultado para la primera derivada es que se llega al de la segunda derivada, además de ser necesario para la prueba del resultado principal. La proposición correspondiente a la primera derivada la enunciamos a continuación, y es adaptada de [16, proposición 5.1], y de [11, sección 4.1] aunque con la notación de este texto, y directamente suponiendo que el término de deriva b en [16] es cero.

Teorema 40. Sea u la solución al problema (1.10). Supongamos además que $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ es de clase $C^1(\mathbb{R}, \mathbb{R})$, y su derivada es una función Lipschitz y acotada. Entonces para cualquier $(t, x) \in [0, T] \times \mathbb{R}$, $u(t, x) \in \bigcap_{p \geq 1} \mathbb{D}^{1,p}$ y además, la derivada de Malliavin cumple que

$$D_{s,y}u(t, x) = p_{t-s}(x - y)\sigma(u(s, y)) + \int_{[s,t] \times \mathbb{R}} p_{t-\tau}(x - \xi)\sigma'(u(\tau, \xi))D_{s,y}u(\tau, \xi)W(d\tau, d\xi), \quad (4.2)$$

para casi cualquier $s \in [0, t]$ y $y \in \mathbb{R}$. Más aún, se cumple que

$$\|D_{s,y}u(t, x)\|_p \leq C_{T,p} p_{t-s}(x - y), \quad (4.3)$$

para cualquier $0 \leq s < t \leq T$, y para $x, y \in \mathbb{R}$.

Con el resultado anterior a la mano, procedemos a enunciar lo respectivo para la segunda derivada de u , tal y como se enuncia en [11]. Es preciso destacar que, de acuerdo con los autores del artículo, la cota de la segunda derivada de la solución mild resuelve un problema abierto en el cálculo de Malliavin y por lo tanto este resultado tiene su interés en sí mismo.

Teorema 41. Sea u la solución al problema (1.10). Supongamos que σ cumple las hipótesis del teorema 39. Fijando $(t, x) \in [0, \infty) \times \mathbb{R}$, tenemos que $u(t, x) \in \bigcap_{p \geq 2} \mathbb{D}^{2,p}$ y para casi todo $0 < r < s < t$, $y, z \in \mathbb{R}$, la segunda derivada de u , que denotamos como $D_{r,z}D_{s,y}u(t, x)$,

satisface la siguiente ecuación diferencial estocástica lineal:

$$\begin{aligned} D_{r,z}D_{s,y}u(t,x) &= p_{t-s}(x-y)\sigma'(u(s,y))D_{r,z}u(s,y) \\ &+ \int_{[s,t]\times\mathbb{R}} p_{t-\tau}(x-\xi)\sigma''(u(\tau,\xi))D_{r,z}u(\tau,\xi)D_{s,y}u(\tau,\xi)W(d\tau,d\xi) \\ &+ \int_{[s,t]\times\mathbb{R}} p_{t-\tau}(x-\xi)\sigma'(u(\tau,\xi))D_{r,z}D_{s,y}u(\tau,\xi)W(d\tau,d\xi) \end{aligned} \quad (4.4)$$

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

$$\|D_{r,z}D_{s,y}u(t,x)\|_p \leq C_{T,p} \Phi_{r,z,s,y}(t,x), \quad (4.5)$$

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

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

Por ser de carácter fundamentalmente técnico, la demostración de este resultado se omite en esta tesis. No obstante, mencionamos que la idea de su prueba consiste en la técnica de las iteraciones de Picard, la desigualdad de Burkholder-Davis-Gundy, y lemas técnicos que se encuentran en [9, lema A.1] y [11, lema A.1, lema A.2]. Es de destacar que las hipótesis del teorema anterior implican las hipótesis del teorema 40.

Tornamos nuestra atención a la existencia de los momentos inversos de la derivada de la solución u al problema del calor (1.10).

Teorema 42. Sea u la solución mild a la ecuación del calor estocástica. 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

$$\sup_{R \geq R_0} \mathbb{E} \left[|D_{v_{R,t}} F_{R,t}|^{-p} \right] < \infty. \quad (4.6)$$

La prueba de este resultado se omite en este texto, por lo que se remite a [11, proposición 4.2] para su consulta. En la prueba del mismo se utilizan nuevamente la desigualdad de Burkholder-Davis-Gundy, propiedades de la solución mild u y distintos lemas presentes en el apéndice del artículo citado.

4.2. Prueba del resultado principal

Con intención de desmenuzar con suficiente claridad la prueba, describimos brevemente su estructura.

La prueba parte de la cota (3.5) aplicada a los promedios espaciales $F_{R,t}$. Es por lo tanto necesario verificar que dichas variables cumplen las hipótesis del teorema 37. Asimismo, serán

usadas constantemente las desigualdades (4.3), (4.5) y (4.6), por lo que debemos verificar que las hipótesis de los teoremas 40, 41 y 42 se satisfacen. Habilitados para usar dichos resultados, de la cota (3.5) nos surgen dos sumandos. El primero de ellos se puede acotar con relativa rapidez apoyándonos de las cotas anteriores y usando la prueba de [9, teorema 1.1], por lo que el grueso de la prueba consiste en acotar el segundo sumando, correspondiente a la norma de la segunda derivada de los promedios espaciales. Allí se descompondrá dicha derivada de segundo orden en dos términos integrales y se procede a acotar cada término por separado utilizando las propiedades de la integral de Itô-Walsh, las cotas mencionadas antes y algunos resultados técnicos presentes en el apéndice de este texto.

Por lo anterior, la prueba está dividida en tres partes. La primera de ellas dedicada a verificar la validez de las hipótesis de los resultados a usar, la segunda de ellas está dirigida a acotar el primero de los sumandos y la última parte se dedica a acotar la norma de la segunda derivada de los promedios espaciales. Dicho todo lo anterior, comenzamos la prueba del resultado principal.

Demostración (del teorema 39).

Parte 1 (Validez de las hipótesis para las cotas). Sean R y $t > 0$ fijos. Consideremos a los promedios espaciales $F_{R,t}$ de la solución mild u al problema (1.10). Buscamos primero ver que $F_{R,t}$ satisface las hipótesis del teorema 37. Para ello, debemos ver que dicha variable puede verse como la divergencia de algún elemento $v \in \mathbb{D}^{1,6}(\mathfrak{H})$. Obsérvese que, de la definición de los promedios espaciales y de la relación en el teorema 14, tenemos las siguientes igualdades.

$$\begin{aligned}
F_{R,t} &= \frac{1}{\sigma_{R,t}} \left(\int_{[-R,R]} u(t,x) dx - 2R \right) \\
&= \frac{1}{\sigma_{R,t}} \left(\int_{[-R,R]} \left(1 + \int_{[0,t] \times \mathbb{R}} p_{t-s}(x-y) \sigma(u(s,y)) W(ds, dy) \right) dx - 2R \right) \\
&= \frac{1}{\sigma_{R,t}} \left(\int_{[-R,R]} \int_{[0,t] \times \mathbb{R}} p_{t-s}(x-y) \sigma(u(s,y)) W(ds, dy) \right) dx \\
&= \int_{[0,t] \times \mathbb{R}} \frac{1}{\sigma_{R,t}} \left(\int_{[-R,R]} p_{t-s}(x-y) \sigma(u(s,y)) dx \right) W(ds, dy) \\
&= \int_{[0,\infty) \times \mathbb{R}} \mathbb{1}_{[0,t]}(s) \frac{1}{\sigma_{R,t}} \left(\int_{[-R,R]} p_{t-s}(x-y) \sigma(u(s,y)) dx \right) W(ds, dy), \tag{4.7}
\end{aligned}$$

en donde en la penúltima igualdad hemos hecho uso del teorema de Fubini estocástico. Por otro lado, sabemos que la integral de Itô-Walsh coincide con el operador de divergencia δ en dos dimensiones, así que definiendo $v_{R,t}$ como

$$v_{R,t} := \mathbb{1}_{[0,t]}(s) \frac{1}{\sigma_{R,t}} \left(\int_{[-R,R]} p_{t-s}(x-y) \sigma(u(s,y)) dx \right), \tag{4.8}$$

la relación (4.7) se convierte en

$$F_{R,t} = \int_{[0,\infty) \times \mathbb{R}} \mathbf{1}_{[0,t]}(s) \frac{1}{\sigma_{R,t}} \left(\int_{[-R,R]} p_{t-s}(x-y) \sigma(u(s,y)) dx \right) W(ds, dy) = \delta(v_{R,t}). \quad (4.9)$$

Por lo tanto, los promedios espaciales se pueden ver como el operador divergencia aplicado a $v_{R,t}$. Observamos ahora que de las tres hipótesis sobre σ que tenemos, estamos en condiciones de aplicar la cota (4.5), por lo que para $(t, x) \in [0, \infty) \times \mathbb{R}$ fijo, $u(t, x) \in \bigcap_{p \geq 2} \mathbb{D}^{2,p}$. Además, para cualquier $T > 0$ y $p \geq 2$, se tiene que

$$\sup_{(t,x) \in [0,\infty) \times \mathbb{R}} \|u(t, x)\|_{2,p} < \infty. \quad (4.10)$$

Esto nos dice que para un $t > 0$ fijo, $F_{R,t} \in \mathbb{D}^{2,6}$ y $v_{R,t} \in \mathbb{D}^{1,2}(\mathfrak{H})$.

Por otro lado, por hipótesis existe algún $q > 10$ tal que $\mathbb{E} [|\sigma(u(t, 0))|^{-q}] < \infty$, de tal forma que eligiendo $p = 4$ en el teorema 42, tenemos que para $q' := 2q > 2(10) = 5(4) = 5p$, al ser el caso que $2q' > q$ y ser Ω un espacio de medida finita,

$$\mathbb{E} [|\sigma(u(t, 0))|^{-2q'}] = \mathbb{E} \left[\left| \frac{1}{\sigma(u(t, 0))} \right|^{2q'} \right] \leq \mathbb{E} \left[\left| \frac{1}{\sigma(u(t, 0))} \right|^q \right] = \mathbb{E} [|\sigma(u(t, 0))|^{-q}] < \infty,$$

por lo que las hipótesis del teorema 42 se satisfacen para $p = 4$ y por lo tanto, para $t > 0$ fijo, existe una constante R_0 tal que la cota de los momentos inversos (4.6) es válida. En particular, para $t > 0$ fijo y $R > R_0$ se sigue que $D_{v_{R,t}} F_{R,t} \in L^4(\Omega)$. Finalmente, por definición estamos tomando los promedios espaciales como variables centradas y con varianza 1.

Hemos mostrado que para $R > R_0$ y $t > 0$, las hipótesis del teorema 37 se cumplen y por lo tanto, la variable $F_{R,t}/D_{v_{R,t}} F_{R,t} \in \text{Dom}(\delta)$, los promedios espaciales tienen una densidad continua y acotada, que denotamos como $f_{F_{R,t}}$, y la misma cumple (3.5). Con precisión, sustituyendo en dicha cota obtenemos que

$$\begin{aligned} \sup_{x \in \mathbb{R}} |f_{F_{R,t}}(x) - \Phi(x)| &\leq \left(\|F_{R,t}\|_4 \| (D_{v_{R,t}} F_{R,t})^{-1} \|_4 + 2 \right) \|1 - D_{v_{R,t}} F_{R,t}\|_2 \\ &\quad + \| (D_{v_{R,t}} F_{R,t})^{-1} \|_4^2 \|D_{v_{R,t}} (D_{v_{R,t}} F_{R,t})\|_2, \end{aligned} \quad (4.11)$$

Observamos dos sumandos en la desigualdad anterior. Mostramos a continuación que cada uno de estos sumandos está acotada por una cantidad proporcional a $1/\sqrt{R}$.

Parte 2 (Cota del primer sumando). Notemos primero que, de la relación (4.11), al ser $F_{R,t} \in \mathbb{D}^{2,6}$, y darse las contenciones $\mathbb{D}^{2,6} \subseteq \mathbb{D}^{2,4} \subseteq L^4(\Omega)$, se sigue que $\|F_{R,t}\|_4 < \infty$. Más aún, dicha cota es uniforme en R gracias a 4.10. Por otro lado, para $t > 0$ y $R > R_0$ fijos, de (4.6) se sigue que

$$\| (D_{v_{R,t}} F_{R,t})^{-1} \|_4 = \mathbb{E} \left[\left| (D_{v_{R,t}} F_{R,t})^{-1} \right|^4 \right]^{1/4} = \mathbb{E} \left[|D_{v_{R,t}} F_{R,t}|^{-4} \right]^{1/4} < \infty, \quad (4.12)$$

y nuevamente gracias a 4.6, dicha cota se puede hacer uniforme en R .

Definiendo $C_t^{(1)} := \left(\|F_{R,t}\|_4 \| (D_{v_{R,t}} F_{R,t})^{-1} \|_4 + 2 \right)$, de lo anterior tenemos que $0 < C_t^{(1)} < \infty$, y dicha constante depende solamente de $t > 0$ fijo. Resta estudiar el término $\|1 - D_{v_{R,t}} F_{R,t}\|_2$. Esta parte de la prueba está basada en la prueba de [9, teorema 1.1]. Dado que $F_{R,t}$ es una variable centrada y de varianza 1, se sigue que

$$\begin{aligned} \|1 - D_{v_{R,t}} F_{R,t}\|_2 &= \mathbb{E} \left[(\mathbb{E} [F_{R,t}^2] - D_{v_{R,t}} F_{R,t})^2 \right]^{1/2} \\ &= \mathbb{E} \left[(\mathbb{E} [F_{R,t} \delta(v_{R,t})] - D_{v_{R,t}} F_{R,t})^2 \right]^{1/2} \\ &= \mathbb{E} \left[(\mathbb{E} [D_{v_{R,t}} F_{R,t}] - D_{v_{R,t}} F_{R,t})^2 \right]^{1/2} \\ &= \sqrt{\text{Var} (D_{v_{R,t}} F_{R,t})}, \end{aligned}$$

donde hemos hecho uso de la definición de δ . Llegados a este punto, remitimos a la prueba del teorema 1.1 en el artículo [9], en donde el grueso de la prueba consiste en acotar $\sqrt{\text{Var} (D_{v_{R,t}} F_{R,t})}$ por C/\sqrt{R} , donde C es una constante que depende de $t > 0$ y de p , que en nuestro caso está fijado en $p = 4$. Por lo tanto, definiendo $C_t^{(2)} := C_t^{(1)} \cdot C$, para $t > 0$,

$$\left(\|F_{R,t}\|_4 \| (D_{v_{R,t}} F_{R,t})^{-1} \|_4 + 2 \right) \|1 - D_{v_{R,t}} F_{R,t}\|_2 \leq \frac{C_t^{(2)}}{\sqrt{R}}, \quad (4.13)$$

donde $C_t^{(2)}$ depende solamente de $t > 0$.

Parte 3 (Cota del segundo sumando). Directamente notamos que el término $\| (D_{v_{R,t}} F_{R,t})^{-1} \|_4^2$ del segundo sumando está acotado, según vimos en (4.12), por lo que nos enfocamos en acotar el término $\|D_{v_{R,t}} (D_{v_{R,t}} F_{R,t})\|_2$. Para tal efecto, establecemos la siguiente notación que nos será útil. Definimos, para $t > 0$ y $R > 0$,

$$Q_R := [-R, R], \quad \Sigma_{t,x} := \sigma(u(t, x)), \quad \Sigma_{t,x}^{(1)} := \sigma'(u(t, x)) \quad \text{y} \quad \Sigma_{t,x}^{(2)} := \sigma''(u(t, x)),$$

y además, para $0 < s < t$, y $y \in \mathbb{R}$, definimos

$$\phi_{R,t}(s, y) := \frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-s}(x - y) dx. \quad (4.14)$$

Armados con la notación anterior, directamente de la definición de los promedios espaciales tenemos que la derivada de Malliavin de los mismos está dada por

$$D_{r,z} F_{R,t} = D_{r,z} \left(\frac{1}{\sigma_{R,t}} \int_{-R,R} u(t, x) dx - 2R \right) = \frac{1}{\sigma_{R,t}} \int_{Q_R} D_{r,z} u(t, x) dx,$$

por lo que tomando la derivada en dirección de $v_{R,t}$ definido en (4.8) y sustituyendo, obtenemos

$$\begin{aligned}
 D_{v_{R,t}} F_{R,t} &= \langle D_{r,z} F_{R,t}, v_{R,t} \rangle_{L^2([0,\infty) \times \mathbb{R})} \\
 &= \int_{[0,\infty)} \int_{\mathbb{R}} v_{R,t} D_{r,z} F_{R,t} dz dr \\
 &= \int_{[0,\infty)} \int_{\mathbb{R}} \left(\mathbb{1}_{[0,t]}(r) \frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-r}(x_1 - z) \sigma(u(r, z)) dx_1 \right) \left(\frac{1}{\sigma_{R,t}} \int_{Q_R} D_{r,z} u(t, x_2) dx_2 \right) dz dr \\
 &= \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{\mathbb{R}} \int_{Q_R} \left(\frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-r}(x_1 - z) dx_1 \right) \sigma(u(r, z)) D_{r,z} u(t, x_2) dx_2 dz dr \\
 &= \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{\mathbb{R}} \int_{Q_R} \phi_{R,t}(r, z) \sigma(u(r, z)) D_{r,z} u(t, x_2) dx_2 dz dr.
 \end{aligned}$$

Haciendo un cambio de variables en esta última igualdad para simplificar notación, llegamos a que

$$D_{v_{R,t}} F_{R,t} = \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{\mathbb{R}} \int_{Q_R} \phi_{R,t}(s, y) \sigma(u(s, y)) D_{s,y} u(t, x) dx dy ds,$$

Tomamos nuevamente la derivada de Malliavin de la expresión anterior para obtener

$$\begin{aligned}
 D_{r,z}(D_{v_{R,t}} F_{R,t}) &= \frac{1}{\sigma_{R,t}} D_{r,z} \left(\int_{[0,t]} \int_{\mathbb{R}} \int_{Q_R} \phi_{R,t}(s, y) \sigma(u(s, y)) D_{s,y} u(t, x) dx dy ds \right) \\
 &= \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{\mathbb{R}} \int_{Q_R} \phi_{R,t}(s, y) D_{r,z} (\sigma(u(s, y)) D_{s,y} u(t, x)) dx dy ds \\
 &= \frac{1}{\sigma_{R,t}} \int_{[r,t]} \int_{\mathbb{R}} \int_{Q_R} \phi_{R,t}(s, y) \sigma'(u(s, y)) D_{r,z} u(s, y) D_{s,y} u(t, x) dx dy ds \\
 &\quad + \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{\mathbb{R}} \int_{Q_R} \phi_{R,t}(s, y) \sigma(u(s, y)) D_{r,z} D_{s,y} u(t, x) dx dy ds,
 \end{aligned}$$

Sacando nuevamente la derivada direccional, la expresión anterior nos lleva a las siguientes

igualdades.

$$\begin{aligned}
D_{v_{R,t}}(D_{v_{R,t}}F_{R,t}) &= \langle D_{r,z}(D_{v_{R,t}}F_{R,t}), v_{R,t} \rangle_{L^2([0,\infty) \times \mathbb{R})} \\
&= \int_{[0,\infty)} \int_{\mathbb{R}} \left(\mathbb{1}_{[0,t]}(r) \frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-r}(x_1 - z) \sigma(u(r, z)) dx_1 \right) \\
&\quad \cdot \left(\frac{1}{\sigma_{R,t}} \int_{[r,t]} \int_{\mathbb{R}} \int_{Q_R} \phi_{R,t}(s, y) \sigma'(u(s, y)) D_{r,z} u(s, y) D_{s,y} u(t, x) dx dy ds \right) dz dr \\
&\quad + \int_{[0,\infty)} \int_{\mathbb{R}} \left(\mathbb{1}_{[0,t]}(r) \frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-r}(x_1 - z) \sigma(u(r, z)) dx_1 \right) \\
&\quad \cdot \left(\frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{\mathbb{R}} \int_{Q_R} \phi_{R,t}(s, y) \sigma(u(s, y)) D_{r,z} D_{s,y} u(t, x) dx dy ds \right) dz dr \\
&= \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{Q_R} \left(\frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-r}(x_1 - z) dx_1 \right) \\
&\quad \cdot (\Sigma_{r,z} \phi_{R,t}(s, y) \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) D_{s,y} u(t, x)) dx dy dz ds dr \\
&\quad + \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{Q_R} \left(\frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-r}(x_1 - z) dx_1 \right) \\
&\quad \cdot (\Sigma_{r,z} \phi_{R,t}(s, y) \Sigma_{s,y} D_{r,z} D_{s,y} u(t, x)) dx dy dz ds dr,
\end{aligned}$$

y reconociendo las respectivas funciones $\phi_{R,t}$, llegamos a que

$$\begin{aligned}
D_{v_{R,t}}(D_{v_{R,t}}F_{R,t}) &= \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{Q_R} \phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) D_{s,y} u(t, x) dx dy dz ds dr \quad (4.15)
\end{aligned}$$

$$\begin{aligned}
&+ \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{Q_R} \phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y} D_{r,z} D_{s,y} u(t, x) dx dy dz ds dr, \quad (4.16)
\end{aligned}$$

donde hemos hecho uso múltiples veces del teorema de Fubini y de la notación previamente establecida. A continuación, utilizamos las fórmulas (4.2) y (4.4) en los dos sumandos de la última igualdad anterior. Así, para el sumando (4.15), sustituyendo la derivada $D_{s,y} u(t, x)$ se tiene la expresión

$$\begin{aligned}
&\frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{Q_R} \phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) D_{s,y} u(t, x) dx dy dz ds dr \\
&= \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{Q_R} (\phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y)) \\
&\quad \cdot \left(p_{t-s}(x - y) \Sigma_{s,y} + \int_{[s,t] \times \mathbb{R}} p_{t-\tau}(x - \xi) \Sigma_{\tau,\xi}^{(1)} D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) dx dy dz ds dr,
\end{aligned}$$

Distribuimos la suma anterior en la integral, por lo que

$$\begin{aligned}
& \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{Q_R} \phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) D_{s,y} u(t, x) dx dy dz ds dr \\
&= \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \left(\frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-s}(x - y) dx \right) \phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) dy dz ds dr \\
&+ \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} (\phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y)) \\
&\cdot \left(\int_{[s,t] \times \mathbb{R}} p_{t-\tau}(x - \xi) \Sigma_{\tau,\xi}^{(1)} D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) dx dy dz ds dr \\
&= \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) dy dz ds dr \\
&+ \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} (\phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y)) \\
&\cdot \left(\int_{[s,t] \times \mathbb{R}} \left(\frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-\tau}(x - \xi) dx \right) \Sigma_{\tau,\xi}^{(1)} D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) dy dz ds dr \\
&= \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) dy dz ds dr \tag{4.17}
\end{aligned}$$

$$\begin{aligned}
&+ \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} (\phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y)) \\
&\cdot \left(\int_{[s,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \Sigma_{\tau,\xi}^{(1)} D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) dy dz ds dr. \tag{4.18}
\end{aligned}$$

Haciendo lo respectivo para el sumando (4.16), obtenemos las siguientes expresiones.

$$\begin{aligned}
& \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{Q_R} \phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y} D_{r,z} D_{s,y} u(t, x) dx dy dz ds dr \\
&= \frac{1}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{Q_R} \phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y} (p_{t-s}(x - y) \Sigma_{s,y}^{(1)} D_{r,z} u(s, y)) \\
&+ \phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y} \left(\int_{[s,t] \times \mathbb{R}} p_{t-\tau}(x - \xi) \Sigma_{\tau,\xi}^{(2)} D_{r,z} u(\tau, \xi) D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) \\
&+ \phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y} \left(\int_{[s,t] \times \mathbb{R}} p_{t-\tau}(x - \xi) \Sigma_{\tau,\xi}^{(1)} D_{r,z} D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) dx dy dz ds dr \\
&= \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) \left(\frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-s}(x - y) dx \right) \\
&+ (\phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}) \\
&\cdot \left(\int_{[s,t] \times \mathbb{R}} \left(\frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-\tau}(x - \xi) dx \right) \Sigma_{\tau,\xi}^{(2)} D_{r,z} u(\tau, \xi) D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) \\
&+ (\phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}) \\
&\cdot \left(\int_{[s,t] \times \mathbb{R}} \left(\frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t-\tau}(x - \xi) dx \right) \Sigma_{\tau,\xi}^{(1)} D_{r,z} D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) dy dz ds dr,
\end{aligned}$$

Y nuevamente reconociendo los términos $\phi_{R,t}$, la expresión anterior es igual a

$$\int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) dy dz ds dr \quad (4.19)$$

$$+ \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} (\phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}) \cdot \left(\int_{[s,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \Sigma_{\tau,\xi}^{(2)} D_{r,z} u(\tau, \xi) D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) dy dz ds dr \quad (4.20)$$

$$+ \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} (\phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}) \cdot \left(\int_{[s,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \Sigma_{\tau,\xi}^{(1)} D_{r,z} D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) dy dz ds dr. \quad (4.21)$$

Observamos que los términos (4.17) y (4.19) coinciden, mientras que el término $\phi_{R,t}(r, z) \phi_{R,t}(s, y)$ es común dentro de las integrales en las expresiones (4.18), (4.20) y (4.21). Luego, definiendo

$$\begin{aligned} Z_{r,z,s,y}(\tau, \xi) &:= \Sigma_{r,z} \Sigma_{s,y}^{(1)} \Sigma_{\tau,\xi}^{(1)} D_{r,z} u(s, y) D_{s,y} u(\tau, \xi) \\ &\quad + \Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(2)} D_{r,z} u(\tau, \xi) D_{s,y} u(\tau, \xi) \\ &\quad + \Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(1)} D_{r,z} D_{s,y} u(\tau, \xi), \end{aligned}$$

y combinando las expresiones (4.17) a (4.21), hallamos que la segunda derivada direccional de $F_{R,t}$ cumple

$$\begin{aligned} D_{v_{R,t}}(D_{v_{R,t}} F_{R,t}) &= 2 \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) dy dz ds dr \\ &\quad + \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} (\phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y)) \\ &\quad \cdot \left(\int_{[s,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \Sigma_{\tau,\xi}^{(1)} D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) dy dz ds dr \\ &\quad + \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} (\phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}) \\ &\quad \cdot \left(\int_{[s,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \Sigma_{\tau,\xi}^{(2)} D_{r,z} u(\tau, \xi) D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) dy dz ds dr \\ &\quad + \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} (\phi_{R,t}(r, z) \phi_{R,t}(s, y) \Sigma_{r,z} \Sigma_{s,y}) \\ &\quad \cdot \left(\int_{[s,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \Sigma_{\tau,\xi}^{(21)} D_{r,z} D_{s,y} u(\tau, \xi) W(d\tau, d\xi) \right) dy dz ds dr, \end{aligned}$$

y denotando por comodidad la primera integral de la suma anterior como $\mathcal{Y}_{R,t}^{(1)}$,

$$\begin{aligned}
& D_{v_{R,t}}(D_{v_{R,t}}F_{R,t}) \\
&= \mathcal{Y}_{R,t}^{(1)} + \int_{[0,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \\
&\quad \cdot \left(\int_{[0,\tau]} \int_{[r,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}(s, y) \left[\Sigma_{r,z} \Sigma_{s,y}^{(1)} \Sigma_{\tau,\xi}^{(1)} D_{r,z} u(s, y) D_{s,y} u(\tau, \xi) \right] dy dz ds dr \right) W(d\tau, d\xi) \\
&\quad + \int_{[0,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \\
&\quad \cdot \left(\int_{[0,\tau]} \int_{[r,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}(s, y) \left[\Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(2)} D_{r,z} u(\tau, \xi) D_{s,y} u(\tau, \xi) \right] dy dz ds dr \right) W(d\tau, d\xi) \\
&\quad + \int_{[0,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \\
&\quad \cdot \left(\int_{[0,\tau]} \int_{[r,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}(s, y) \left[\Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(1)} D_{r,z} D_{s,y} u(\tau, \xi) \right] dy dz ds dr \right) W(d\tau, d\xi) \\
&= \mathcal{Y}_{R,t}^{(1)} + \int_{[0,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \left(\int_{[0,\tau]} \int_{[r,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}(s, y) Z_{r,z,s,y}(\tau, \xi) dy dz ds dr \right) W(d\tau, d\xi) \\
&= 2\mathcal{Y}_{R,t}^{(1)} + \mathcal{Y}_{R,t}^{(2)}, \tag{4.22}
\end{aligned}$$

en donde convenimos que

$$\begin{aligned}
\mathcal{Y}_{R,t}^{(1)} &:= \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) dy dz ds dr \\
\mathcal{Y}_{R,t}^{(2)} &:= \int_{[0,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \left(\int_{[0,\tau]} \int_{[r,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}(s, y) Z_{r,z,s,y}(\tau, \xi) dy dz ds dr \right) W(d\tau, d\xi),
\end{aligned}$$

y donde nuevamente se ha usado múltiples veces del teorema de Fubini y la notación previamente establecida. Una vez que llegamos a este punto, buscamos acotar cada uno de los sumandos de (4.22).

Comenzamos con el término $\mathcal{Y}_{R,t}^{(1)}$. Primero hallamos una cota para la norma $L^2(\Omega)$ de la expresión $\Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y)$. Usando la desigualdad de Cauchy-Schwarz,

$$\left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) \right\|_2 \leq \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} \right\|_4 \left\| D_{r,z} u(s, y) \right\|_4,$$

y usando ahora la desigualdad (4.3), para $0 \leq r < s \leq T$, $y, z \in \mathbb{R}$ y $p = 4$, hallamos que

$$\left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} \right\|_4 \left\| D_{r,z} u(s, y) \right\|_4 \leq \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} \right\|_4 C_T^{(3)} p_{s-r}(y - z), \tag{4.23}$$

donde $C_T^{(3)}$ es una constante que solo depende de T . Por otro lado, por hipótesis σ' es una función acotada, digamos, por una constante $K_1 > 0$, de forma que

$$\left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} \right\|_4 \leq K_1 \left\| \Sigma_{r,z} \Sigma_{s,y} \right\|_4 \leq K_1 \left\| \Sigma_{r,z} \right\|_8 \left\| \Sigma_{s,y} \right\|_8. \tag{4.24}$$

Ahora bien, como la condición inicial en el problema del calor que estamos tratando es $u_0(w) = u(0, w) = 1$, para cualquier $w \in \mathbb{R}$, y dado que σ es Lipschitz, se tienen las siguientes relaciones para $\Sigma_{r,z}$:

$$\begin{aligned}
\|\Sigma_{r,z}\|_8 &\leq \|\sigma(u(r, z)) - \sigma(u(0, z))\|_8 + \|\sigma(u(0, z))\|_8 \\
&\leq K_2 \|u(r, z) - u(0, z)\| + \|\sigma(1)\|_8 \\
&\leq K_2 (\|u(r, z)\|_8 + \|1\|_8) + |\sigma(1)| \\
&\leq K_2 (C_T^{(4)} + 1) + |\sigma(1)|,
\end{aligned} \tag{4.25}$$

en donde $K_2 > 0$ es la constante de Lipschitz de σ , y en la cuarta desigualdad hemos hecho uso de la igualdad (1.13) con $T > 0$ y $p = 8$. Obsérvese que la cota obtenida antes depende solamente de σ y de $T > 0$. De manera análoga se puede acotar el término $\|\Sigma_{s,y}\|_8$. Por lo tanto, combinando (4.24), (4.25) y la cota (4.25) análoga para $\Sigma_{s,y}$, hallamos que

$$\|\Sigma_{r,z}\Sigma_{s,y}\Sigma_{s,y}^{(1)}\|_4 \leq C_T^{(5)}, \tag{4.26}$$

en donde $C_T^{(5)} > 0$ es una constante que solo depende de $T > 0$ y σ . Insertando esta cota en (4.23), obtenemos que

$$\|\Sigma_{r,z}\Sigma_{s,y}\Sigma_{s,y}^{(1)}\|_4 \|D_{r,z}u(s, y)\|_4 \leq C_T^{(6)} p_{s-r}(y - z), \tag{4.27}$$

donde $C_T^{(6)} := C_T^{(5)} \cdot C_T^{(3)}$. Obtenida esta cota, procedemos a extraer la norma $L^2(\Omega)$ de $\mathcal{Y}_{R,t}^{(1)}$. Obtenemos así las siguientes relaciones.

$$\begin{aligned}
\|\mathcal{Y}_{R,t}^{(1)}\|_2 &= \left(\int_{\Omega} \left(\int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) dy dz ds dr \right)^2 d\mathbb{P} \right)^{1/2} \\
&= \left(\int_{\Omega} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{[0,t]} \int_{[r',t]} \int_{\mathbb{R}^2} (\phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y)) \right. \\
&\quad \cdot \left(\phi_{R,t}(r', z') \phi_{R,t}^2(s', y') \Sigma_{r',z'} \Sigma_{s',y'} \Sigma_{s',y'}^{(1)} D_{r',z'} u(s', y') \right) dy dz ds dr dy' dz' ds' dr' d\mathbb{P} \Big)^{1/2} \\
&= \left(\int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{[0,t]} \int_{[r',t]} \int_{\mathbb{R}^2} (\phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \phi_{R,t}(r', z') \phi_{R,t}^2(s', y')) \right. \\
&\quad \cdot \left(\int_{\Omega} \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) \Sigma_{r',z'} \Sigma_{s',y'} \Sigma_{s',y'}^{(1)} D_{r',z'} u(s', y') \right) dy dz ds dr dy' dz' ds' dr' d\mathbb{P} \Big)^{1/2}.
\end{aligned}$$

Usando ahora la desigualdad de Cauchy-Schwarz,

$$\begin{aligned}
 \left\| \mathcal{Y}_{R,t}^{(1)} \right\|_2 &\leq \left(\int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \int_{[0,t]} \int_{[r',t]} \int_{\mathbb{R}^2} (\phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \phi_{R,t}(r', z') \phi_{R,t}^2(s', y')) \right. \\
 &\quad \cdot \left(\left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) \right\|_2 \left\| \Sigma_{r',z'} \Sigma_{s',y'} \Sigma_{s',y'}^{(1)} D_{r',z'} u(s', y') \right\|_2 \right) dy dz ds dr dy' dz' ds' dr' \Big)^{1/2} \\
 &= \left(\int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) \right\|_2 dy dz ds dr \right)^{1/2} \\
 &\quad \cdot \left(\int_{[0,t]} \int_{[r',t]} \int_{\mathbb{R}^2} \phi_{R,t}(r', z') \phi_{R,t}^2(s', y') \left\| \Sigma_{r',z'} \Sigma_{s',y'} \Sigma_{s',y'}^{(1)} D_{r',z'} u(s', y') \right\|_2 dy' dz' ds' dr' \right)^{1/2} \\
 &= \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) \right\|_2 dy dz ds dr.
 \end{aligned}$$

Usamos ahora (4.27) en la última desigualdad anterior para obtener que

$$\begin{aligned}
 \left\| \mathcal{Y}_{R,t}^{(1)} \right\|_2 &\leq \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}^2(s, y) \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{s,y}^{(1)} D_{r,z} u(s, y) \right\|_2 dy dz ds dr \\
 &\leq C_T^{(6)} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}^2(s, y) p_{s-r}(y - z) dy dz ds dr.
 \end{aligned}$$

En esta última expresión integramos sobre z y expandimos el término $\phi_{R,t}(r, z)$ para hallar que

$$\begin{aligned}
 \left\| \mathcal{Y}_{R,t}^{(1)} \right\|_2 &\leq C_T^{(6)} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}} \phi_{R,t}^2(s, y) \left(\int_{\mathbb{R}} \phi_{R,t}(r, z) p_{s-r}(z - y) dz \right) dy ds dr \\
 &= C_T^{(6)} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}} \phi_{R,t}^2(s, y) \left(\frac{1}{\sigma_{R,t}} \int_{Q_R} \int_{\mathbb{R}} p_{t-r}(x - z) p_{s-r}(z - y) dz dx \right) dy ds dr \\
 &= C_T^{(6)} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}} \phi_{R,t}^2(s, y) \left(\frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t+s-2r}(x - y) dx \right) dy ds dr,
 \end{aligned}$$

donde hemos hecho uso de la propiedad de semigrupo del núcleo del calor $(p_t)_{t \geq 0}$ en la última igualdad. Claramente se tiene que

$$\frac{1}{\sigma_{R,t}} \int_{Q_R} p_{t+s-2r}(x - y) dx \leq \frac{1}{\sigma_{R,t}} \int_{\mathbb{R}} p_{t+s-2r}(x - y) dx \leq \frac{1}{\sigma_{R,t}},$$

por lo que insertando esta última desigualdad en la anterior, llegamos a que

$$\left\| \mathcal{Y}_{R,t}^{(1)} \right\|_2 \leq \frac{C_T^{(6)}}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}} \phi_{R,t}^2(s, y) dy ds dr.$$

Utilizando el lema A.3, obtenemos que existe un $R_0 \geq 1$ dependiente de $t > 0$ tal que para cualquier $R \geq R_0$,

$$\frac{C_T^{(6)}}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} \int_{\mathbb{R}} \phi_{R,t}^2(s, y) dy ds dr \leq \frac{C_T^{(6)}}{\sigma_{R,t}} \int_{[0,t]} \int_{[r,t]} C_t ds dr \leq \frac{C_t^{(7)}}{\sigma_{R,t}},$$

por lo que tenemos la siguiente cota que es uniforme en R ,

$$\left\| \mathcal{Y}_{R,t}^{(1)} \right\|_2 \leq \frac{C_t^{(7)}}{\sigma_{R,t}}.$$

Finalmente, utilizamos el lema A.2, para argumentar que, como $\frac{\sigma_{R,t}^2}{R} \xrightarrow[n \rightarrow \infty]{} 2 \int_0^t \xi(s) ds$, donde $\xi(s) = \mathbb{E}[\sigma^2(u(s, y))]$, y esta última función no depende de $y \in \mathbb{R}$ (ver el enunciado del lema mencionado) entonces existe un $R_1 > 0$ tal que para cualquier $R > R_1$,

$$\left\| \mathcal{Y}_{R,t}^{(1)} \right\|_2 \leq \frac{C_t^{(7)}}{\sigma_{R,t}} = \frac{C_t^{(7)}}{\sqrt{\frac{\sigma_{R,t}^2}{R}} \sqrt{R}} \leq \frac{C_t^{(8)}}{\sqrt{R}}, \quad (4.28)$$

en donde $C_t^{(8)}$ es una constante que depende de $C_t^{(7)}$ y de la integral en $[0, t]$ de ξ , función que es independiente de y , y por ende tanto la integral como la constante $C_t^{(8)}$ sólo dependen de t . Es menester comentar que para el uso del lema A.2 y garantizar la validez de la cota anterior, la función σ debe ser cumplir que la esperanza de la misma evaluada en u sea distinta de cero. Dicha condición bien puede ser exigida como parte de las hipótesis del teorema.

Procedemos ahora a estimar el término $\mathcal{Y}_{R,t}^{(2)}$. Extrayendo el cuadrado de la norma $L^2(\Omega)$ de dicho término, y utilizando la isometría de Itô-Walsh (análoga a la isometría de Itô en cálculo estocástico usual), tenemos las siguientes igualdades.

$$\begin{aligned} \left\| \mathcal{Y}_{R,t}^{(2)} \right\|_2^2 &= \left\| \int_{[0,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \left(\int_{[0,\tau]} \int_{[r,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}(s, y) Z_{r,z,s,y}(\tau, \xi) dy dz ds dr \right) W(d\tau, d\xi) \right\|_2^2 \\ &= \mathbb{E} \left[\left(\int_{[0,t] \times \mathbb{R}} \phi_{R,t}(\tau, \xi) \left(\int_{[0,\tau]} \int_{[r,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}(s, y) Z_{r,z,s,y}(\tau, \xi) dy dz ds dr \right) W(d\tau, d\xi) \right)^2 \right] \\ &= \mathbb{E} \left[\int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \left(\int_{[0,\tau]} \int_{[r,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}(s, y) Z_{r,z,s,y}(\tau, \xi) dy dz ds dr \right)^2 d\xi d\tau \right] \\ &= \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \mathbb{E} \left[\left(\int_{[0,\tau]} \int_{[r,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}(s, y) Z_{r,z,s,y}(\tau, \xi) dy dz ds dr \right)^2 \right] d\xi d\tau. \end{aligned}$$

Reescribiendo el cuadrado de la integral anterior, tenemos que

$$\begin{aligned} \left\| \mathcal{Y}_{R,t}^{(2)} \right\|_2^2 &= \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \mathbb{E} \left[\left(\int_{[0,\tau]} \int_{[r,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r, z) \phi_{R,t}(s, y) Z_{r,z,s,y}(\tau, \xi) dy dz ds dr \right)^2 \right] d\xi d\tau \\ &= \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \mathbb{E} \left[\left(\int_{[0,\tau]} \int_{[r_1,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r_1, z_1) \phi_{R,t}(s_1, y_1) Z_{r_1,z_1,s_1,y_1}(\tau, \xi) dy_1 dz_1 ds_1 dr_1 \right) \right. \\ &\quad \cdot \left. \left(\int_{[0,\tau]} \int_{[r_2,\tau]} \int_{\mathbb{R}^2} \phi_{R,t}(r_2, z_2) \phi_{R,t}(s_2, y_2) Z_{r_2,z_2,s_2,y_2}(\tau, \xi) dy_2 dz_2 ds_2 dr_2 \right) \right] d\xi d\tau \\ &= \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \int_{0 \leq r_1 \leq s_1 \leq \tau} \int_{0 \leq r_2 \leq s_2 \leq \tau} \int_{\mathbb{R}^4} \phi_{R,t}(r_1, z_1) \phi_{R,t}(s_1, y_1) \phi_{R,t}(r_2, z_2) \phi_{R,t}(s_2, y_2) \\ &\quad \cdot \mathbb{E} [Z_{r_1,z_1,s_1,y_1}(\tau, \xi) Z_{r_2,z_2,s_2,y_2}(\tau, \xi)] dy_1 dz_1 ds_1 dr_1 dy_2 dz_2 ds_2 dr_2 d\xi d\tau, \end{aligned}$$

en donde $Z_{r_1, z_1, s_1, y_1}(\xi, \tau)$ y $Z_{r_2, z_2, s_2, y_2}(\xi, \tau)$ son copias i.i.d. de $Z_{r, z, s, y}(\xi, \tau)$. A continuación, utilizamos la desigualdad de Cauchy-Schwarz justo en estos términos para obtener que

$$\begin{aligned} \|\mathcal{Y}_{R,t}^{(2)}\|_2^2 &\leq \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \int_{0 \leq r_1 \leq s_1 \leq \tau} \int_{0 \leq r_2 \leq s_2 \leq \tau} \int_{\mathbb{R}^4} \phi_{R,t}(r_1, z_1) \phi_{R,t}(s_1, y_1) \phi_{R,t}(r_2, z_2) \phi_{R,t}(s_2, y_2) \\ &\quad \cdot \|Z_{r_1, z_1, s_1, y_1}(\tau, \xi)\|_2 \|Z_{r_2, z_2, s_2, y_2}(\tau, \xi)\|_2 dy_1 dz_1 ds_1 dr_1 dy_2 dz_2 ds_2 dr_2 d\xi d\tau \\ &= \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \int_{0 \leq r_1 \leq s_1 \leq \tau} \int_{0 \leq r_2 \leq s_2 \leq \tau} \int_{\mathbb{R}^4} \prod_{i=1,2} \phi_{R,t}(r_i, z_i) \phi_{R,t}(s_i, y_i) \\ &\quad \cdot \|Z_{r_i, z_i, s_i, y_i}(\tau, \xi)\|_2 dy_1 dz_1 ds_1 dr_1 dy_2 dz_2 ds_2 dr_2 d\xi d\tau \end{aligned} \quad (4.29)$$

En este punto, nos interesamos por acotar el término $\|Z_{r,z,s,y}(\xi, \tau)\|_2$. Usando la definición de $Z_{r,z,s,y}$, la desigualdad de Minkowski, y la desigualdad de Hölder,

$$\begin{aligned} &\|Z_{r,z,s,y}(\xi, \tau)\|_2 \\ &\leq \left\| \Sigma_{r,z} \Sigma_{s,y}^{(1)} \Sigma_{\tau,\xi}^{(1)} D_{r,z} u(s, y) D_{s,y} u(\tau, \xi) \right\|_2 + \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(2)} D_{r,z} u(\tau, \xi) D_{s,y} u(\tau, \xi) \right\|_2 \\ &\quad + \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(1)} D_{r,z} D_{s,y} u(\tau, \xi) \right\|_2 \\ &\leq \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(1)} \right\|_4 \|D_{r,z} u(s, y)\|_8 \|D_{s,y} u(\tau, \xi)\|_8 + \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(2)} \right\|_4 \|D_{r,z} u(\tau, \xi)\|_8 \|D_{s,y} u(\tau, \xi)\|_8 \\ &\quad + \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(1)} \right\|_4 \|D_{r,z} D_{s,y} u(\tau, \xi)\|_4. \end{aligned}$$

En esta última desigualdad utilizamos cuatro veces la relación (4.3) con $p = 8$ y la desigualdad (4.5) con $p = 4$ para obtener que

$$\begin{aligned} \|Z_{r,z,s,y}(\xi, \tau)\|_2 &\leq C_t^{(9)} \left(\left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(1)} \right\|_4 p_{s-r}(y-z) p_{\tau-s}(\xi-y) \right. \\ &\quad \left. + \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(2)} \right\|_4 p_{\tau-r}(\xi-z) p_{\tau-s}(\xi-y) + \left\| \Sigma_{r,z} \Sigma_{s,y} \Sigma_{\tau,\xi}^{(1)} \right\|_4 \phi_{r,z,s,y}(\tau, \xi) \right), \end{aligned}$$

en donde $C_t^{(9)}$ es una constante que solo depende de $t > 0$ y de σ . Ahora bien, por hipótesis, $|\sigma'| \leq K_1$ en concordancia con la notación de la desigualdad (4.24). Por lo tanto si en la última cota usamos la desigualdad de Hölder, así como las desigualdades (4.25) y (4.26), hallamos que

$$\begin{aligned} \|Z_{r,z,s,y}(\xi, \tau)\|_2 &\leq C_t^{(9)} \left(C_t^{(10)} p_{s-r}(y-z) p_{\tau-s}(\xi-y) \right. \\ &\quad \left. + C_t^{(11)} \left\| \Sigma_{\tau,\xi}^{(2)} \right\|_8 p_{\tau-r}(\xi-z) p_{\tau-s}(\xi-y) + C_t^{(12)} \phi_{r,z,s,y}(\tau, \xi) \right). \end{aligned}$$

Finalmente, recordemos que por hipótesis $|\Sigma_{\tau,\xi}^{(2)}| = |\sigma''(u(\tau, \xi))| \leq C(1 + |u(\tau, \xi)|)^m$, para alguna $m > 0$, y $C > 0$. Usando esto y la igualdad (1.13) para $p = 8$ deducimos que

$$\left\| \Sigma_{\tau,\xi}^{(2)} \right\|_8 \leq C \|1 + |u(\tau, \xi)|\|_8 \leq C(1 + \mathbb{E}[|u(\tau, \xi)|^8])^{1/8} = C_t^{(13)},$$

donde $C_t^{(13)}$ es una cota que solo depende de $t > 0$ y σ . Insertando esta cota en la desigualdad previa, hallamos finalmente que

$$\begin{aligned} \|Z_{r,z,s,y}(\xi, \tau)\|_2 &\leq C_t^{(9)} \left(C_t^{(10)} p_{s-r}(y-z) p_{\tau-s}(\xi-y) \right. \\ &\quad \left. + C_t^{(13)} p_{\tau-r}(\xi-z) p_{\tau-s}(\xi-y) + C_t^{(12)} \phi_{r,z,s,y}(\tau, \xi) \right). \end{aligned} \quad (4.30)$$

Ahora, si insertamos la cota (4.30) en (4.29) llegamos a que

$$\begin{aligned} \|\mathcal{Y}_{R,t}^{(2)}\|_2^2 &\leq C_t^{(9)} \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \int_{0 \leq r_1 \leq s_1 \leq \tau} \int_{0 \leq r_2 \leq s_2 \leq \tau} \int_{\mathbb{R}^4} \prod_{i=1,2} \phi_{R,t}(r_i, z_i) \phi_{R,t}(s_i, y_i) \\ &\quad \cdot \left(C_t^{(10)} p_{s_i-r_i}(y_i - z_i) p_{\tau-s_i}(\xi - y_i) + C_t^{(13)} p_{\tau-r_i}(\xi - z_i) p_{\tau-s_i}(\xi - y_i) \right. \\ &\quad \left. + C_t^{(12)} \phi_{r_i, z_i, s_i, y_i}(\tau, \xi) \right) dy_1 dz_1 ds_1 dr_1 dy_2 dz_2 ds_2 dr_2 d\xi d\tau. \end{aligned}$$

Utilizando nuevamente el lema A.3, para $R > R_0 \geq 1$ y $0 < s_i, r_i < t$,

$$\begin{aligned} &\|\mathcal{Y}_{R,t}^{(2)}\|_2^2 \\ &\leq \frac{C_t^{(14)}}{\sigma_{R,t}^2} \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \int_{0 \leq r_1 \leq s_1 \leq \tau} \int_{0 \leq r_2 \leq s_2 \leq \tau} \int_{\mathbb{R}^4} \prod_{i=1,2} \left(C_t^{(10)} p_{s_i-r_i}(y_i - z_i) p_{\tau-s_i}(\xi - y_i) \right. \\ &\quad \left. + C_t^{(13)} p_{\tau-r_i}(\xi - z_i) p_{\tau-s_i}(\xi - y_i) \right. \\ &\quad \left. + C_t^{(12)} \phi_{r_i, z_i, s_i, y_i}(\tau, \xi) \right) dy_1 dz_1 ds_1 dr_1 dy_2 dz_2 ds_2 dr_2 d\xi d\tau. \end{aligned} \quad (4.31)$$

Aquí separamos primero la integral en \mathbb{R}^4 del producto como dos integrales separadas,

$$\begin{aligned} &\int_{\mathbb{R}^4} \prod_{i=1,2} \left(C_t^{(10)} p_{s_i-r_i}(y_i - z_i) p_{\tau-s_i}(\xi - y_i) \right. \\ &\quad \left. + C_t^{(13)} p_{\tau-r_i}(\xi - z_i) p_{\tau-s_i}(\xi - y_i) + C_t^{(12)} \phi_{r_i, z_i, s_i, y_i}(\tau, \xi) \right) dy_1 dz_1 dy_2 dz_2 \\ &= \left(\int_{\mathbb{R}^2} C_t^{(10)} p_{s_1-r_1}(y_1 - z_1) p_{\tau-s_1}(\xi - y_1) + C_t^{(13)} p_{\tau-r_1}(\xi - z_1) p_{\tau-s_1}(\xi - y_1) \right. \\ &\quad \left. + C_t^{(12)} \phi_{r_1, z_1, s_1, y_1}(\tau, \xi) dy_1 dz_1 \right) \left(\int_{\mathbb{R}^2} C_t^{(10)} p_{s_2-r_2}(y_2 - z_2) p_{\tau-s_2}(\xi - y_2) \right. \\ &\quad \left. + C_t^{(13)} p_{\tau-r_2}(\xi - z_2) p_{\tau-s_2}(\xi - y_2) + C_t^{(12)} \phi_{r_2, z_2, s_2, y_2}(\tau, \xi) dy_2 dz_2 \right), \end{aligned}$$

y ahora integramos primero los núcleos gaussianos $p_t(x)$ en las variables y_1, y_2, z_1 y z_2 , y pos-

teriormente utilizamos el lema A.1, hallando que la expresión anterior es igual a

$$\begin{aligned}
&= \left(C_t^{(10)} \int_{\mathbb{R}} p_{s_1-r_1}(y_1 - z_1) \int_{\mathbb{R}} p_{\tau-s_1}(\xi - y_1) dy_1 dz_1 \right. \\
&\quad + C_t^{(13)} \int_{\mathbb{R}} p_{\tau-r_1}(\xi - z_1) \int_{\mathbb{R}} p_{\tau-s_1}(\xi - y_1) dy_1 dz_1 + C_t^{(12)} \int_{\mathbb{R}^2} \phi_{r_1, z_1, s_1, y_1}(\tau, \xi) dy_1 dz_1 \Big) \\
&\quad \cdot \left(C_t^{(10)} \int_{\mathbb{R}} p_{s_2-r_2}(y_2 - z_2) \int_{\mathbb{R}} p_{\tau-s_2}(\xi - y_2) dy_2 dz_2 \right. \\
&\quad + C_t^{(13)} \int_{\mathbb{R}} p_{\tau-r_2}(\xi - z_2) \int_{\mathbb{R}} p_{\tau-s_2}(\xi - y_2) dy_2 dz_2 + C_t^{(12)} \int_{\mathbb{R}^2} \phi_{r_2, z_2, s_2, y_2}(\tau, \xi) dy_2 dz_2 \Big) \\
&= \left(C_t^{(10)} + C_t^{(13)} + \int_{\mathbb{R}^2} \phi_{r_1, z_1, s_1, y_1}(\tau, \xi) dy_1 dz_1 \right) \left(C_t^{(10)} + C_t^{(13)} + \int_{\mathbb{R}^2} \phi_{r_2, z_2, s_2, y_2}(\tau, \xi) dy_2 dz_2 \right) \\
&\leq \left(C_t^{(15)} + C_t^{(16)} \left(1 + \frac{1}{(s_1 - r_1)^{1/4}} \right) \right) \left(C_t^{(15)} + C_t^{(16)} \left(1 + \frac{1}{(s_2 - r_2)^{1/4}} \right) \right), \tag{4.32}
\end{aligned}$$

Combinando las desigualdades (4.31) y (4.32) se sigue que

$$\begin{aligned}
\left\| \mathcal{Y}_{R,t}^{(2)} \right\|_2^2 &\leq \frac{C_t^{(14)}}{\sigma_{R,t}^2} \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \int_{0 \leq r_1 \leq s_1 \leq \tau} \int_{0 \leq r_2 \leq s_2 \leq \tau} \left(C_t^{(15)} + C_t^{(16)} \left(1 + \frac{1}{(s_1 - r_1)^{1/4}} \right) \right) \\
&\quad \cdot \left(C_t^{(15)} + C_t^{(16)} \left(1 + \frac{1}{(s_2 - r_2)^{1/4}} \right) \right) ds_1 dr_1 ds_2 dr_2 d\xi d\tau \\
&= \frac{C_t^{(14)}}{\sigma_{R,t}^2} \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \left(\int_{0 \leq r \leq s \leq \tau} \left(C_t^{(15)} + C_t^{(16)} \left(1 + \frac{1}{(s - r)^{1/4}} \right) \right) ds dr \right)^2 d\xi d\tau \\
&= \frac{C_t^{(17)}}{\sigma_{R,t}^2} \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) \left(\int_{0 \leq r \leq s \leq \tau} \left(1 + \frac{C_t^{(18)}}{(s - r)^{1/4}} \right) ds dr \right)^2 d\xi d\tau.
\end{aligned}$$

Al realizar la integral en s y r , obtenemos que

$$\left(\int_{0 \leq r \leq s \leq \tau} \left(1 + \frac{C_t^{(18)}}{(s - r)^{1/4}} \right) ds dr \right)^2 = \left(\frac{\tau^2}{2} - C_t^{(19)} \tau^{7/4} \right)^2,$$

donde $C^{(19)} > 0$ es una constante que solamente depende de t y de σ . Pero dado que estamos integrando a τ en el intervalo $[0, t]$, el resultado de la integral anterior lo podemos acotar por una constante $C_t^{(20)}$ que depende solamente de $t > 0$ y de σ . Deducimos así que

$$\left\| \mathcal{Y}_{R,t}^{(2)} \right\|_2^2 \leq \frac{C_t^{(20)}}{\sigma_{R,t}^2} \int_{[0,t]} \int_{\mathbb{R}} \phi_{R,t}^2(\tau, \xi) d\xi d\tau.$$

Aplicamos ahora el lema A.3 e integramos sobre τ para obtener que

$$\left\| \mathcal{Y}_{R,t}^{(2)} \right\|_2^2 \leq \frac{C_t^{(21)}}{\sigma_{R,t}^2}.$$

Por último, extraemos raíz y aplicamos nuevamente el lema A.2 para deducir que

$$\left\| \mathcal{Y}_{R,t}^{(2)} \right\|_2^2 \leq \frac{C_t^{(22)}}{\sqrt{R}}, \quad (4.33)$$

en donde $C_t^{(22)}$ depende solamente de $t > 0$ y σ . Combinando las desigualdades (4.13), (4.28) y (4.33) hallamos que

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

donde C_t es una constante positiva que depende solo de $t > 0$ y de σ , completando la prueba. ■

Capítulo 5

Conclusiones y trabajo futuro

El trabajo presente tuvo como objetivo estudiar brevemente los tópicos correspondientes a ecuaciones diferenciales parciales estocásticas, cálculo de Malliavin, método de Stein y existencia de densidades. Todo esto con la finalidad de presentar y demostrar el teorema 39.

Como fue mencionado al inicio, siempre que fue posible las fuentes de las ideas se citaron al inicio de cada capítulo, sección o resultado. La contribución hecha reside principalmente en la búsqueda de una exposición amigable de estos tópicos. Destaca también el intento de «desmenuzar» la prueba del teorema principal para hacerlo más accesible, tratando de reunir el material necesario para el entendimiento y la realización de la prueba de dicho teorema.

Lo anterior se materializó en realizar con más detalle las justificaciones a algunas relaciones o cálculos que involucran a las integrales estocásticas y de Lebesgue, así como a los argumentos de algunos teoremas. Como ejemplos particulares de lo anterior hallamos la prueba del teorema 37, la cual fue realizada con bastante detalle, o bien el uso del lema A.2 en la prueba del teorema principal, en el cual se debe destacar la importancia de que σ sea una función tal que su esperanza sea distinta de cero \mathbb{P} casi seguramente para garantizar que el límite esté bien definido y por lo tanto estar seguros de la validez de la cota (4.28).

A pesar de lo anterior, y como fue dicho a lo largo del texto en múltiples ocasiones, un estudio exhaustivo de cada uno de los tópicos involucrados en este trabajo estuvo siempre fuera del objetivo del mismo por motivos tanto temporales como de la naturaleza propia de esta tesis, por lo que aspectos particulares como la prueba de los lemas del apéndice o de los lemas de la sección 3.2, la prueba de la existencia de los momentos inversos (4.6), o la cota (4.5) han quedado fuera.

Concluimos este último capítulo, y el cuerpo principal de este texto reuniendo algunas ideas en cuanto a lo que respecta a trabajo en otras direcciones, así como trabajo futuro del tópico tratado. Dado que existen múltiples posibilidades, a continuación mencionamos algunas de ellas.

- **Superconvergencia.** Una vez que ha sido establecida la convergencia uniforme de las densidades de los promedios espaciales $f_{F_{R,t}}$ a la densidad de una normal Φ , podemos preguntarnos ¿qué sucede con la convergencia de sus derivadas? ¿Se puede decir algo sobre la regularidad de las densidades? Esta es una pregunta que ha sido planteada en otros tópicos partiendo del fenómeno de *superconvergencia*, visto por primera vez en el teorema central del límite para probabilidad libre en 1995.

Usando las herramientas provistas por el cálculo de Malliavin, este problema bien podría ser abordado a partir del trabajo hecho en este texto, por lo que a propósito es ubicado como primer posible trabajo a futuro.

- **Expansiones de Edgeworth.** Entrados en el estudio de la regularidad de las densidades $f_{F_{R,t}}$, podemos preguntarnos por la posibilidad de realizar una expansión de Edgeworth de la distribución normal, de tal forma que la conclusión del teorema principal, a saber, la cota

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

pueda ser refinada en términos de R , añadiendo *términos correctivos* en la distancia uniforme de las densidades. Normalmente bajo hipótesis de regularidad de las densidades es que es posible añadir dichos términos correctivos.

- **La ecuación estocástica de onda.** Porsupuesto que, entre las ecuaciones diferenciales parciales famosas, se encuentra la ecuación de onda. Es natural preguntarse por lo que sucede en este escenario. Es de destacar que, por mucho que las ecuaciones «se parezcan», el comportamiento de las soluciones clásicas a ambas ecuaciones es bastante distinto, por lo que se podría esperar algo similar en el caso estocástico. Existe trabajo en esta dirección, tal y como se puede ver en [17], donde los autores lograron establecer una tasa de convergencia de los promedios espaciales de la solución a la ecuación de onda estocástica perturbada con un ruido blanco, a la distribución normal, utilizando la distancia de Wasserstein.

Restaría estudiar qué sucede con la existencia de las densidades de los promedios espaciales, y su comportamiento asintótico.

- **Otras versiones de la ecuación.** Si bien la ecuación de onda estocástica es la primera ecuación a la cual tornamos nuestra atención al intentar trasladar los resultados de este texto, otras ecuaciones o bien otras condiciones resultan interesantes de explorar.

Tal es el caso del modelo *Parabolic Anderson* con condición inicial $u_0(x) = \delta_0(x)$, de la ecuación del calor con laplaciano faccionario o de la ecuación del calor estocástica en un dominio distinto y que obligue a imponer condiciones de frontera que bien podrían alterar la dinámica estocástica de los promedios espaciales a largo plazo (por ejemplo, la ecuación del calor en $[0, \infty) \times [0, \infty)$), entre otros. En la dirección del modelo *Parabolic Anderson*, el mismo artículo de Kuzgun y Nualart [11] lo trabaja. En cuanto a la ecuación del

calor con laplaciano fraccionario se puede encontrar en [19] trabajo en el cual los autores, basándose precisamente en el artículo de Kuzgun y Nualart mencionado, logran establecer resultados análogos al teorema principal en tres casos distintos.

- **Ley de grandes números y grandes desviaciones.** Otra dirección en la cual se puede reflexionar, es el estudio de los otros conjuntos de resultados clásicos de probabilidad: la ley de grandes números y grandes desviaciones, pero adaptado al contexto de los promedios espaciales de la ecuación del calor estocástica. En este sentido, el trabajo hecho en [5] muestra que las distribuciones de los promedios espaciales de las soluciones a la ecuación del calor estocástica del calor y de onda cumplen el principio de grandes desviaciones conforme el intervalo donde se promedia tiende a infinito.
- **Otros ruidos, procesos gaussianos más generales.** En el texto presente tratamos una ecuación del calor que estuvo perturbada por un ruido blanco. Sin embargo, podríamos pensar en un ruido coloreado como es el caso del artículo de Kuzgun y Nualart [12], el cual de hecho trata la convergencia de las densidades de los promedios espaciales en el caso del modelo Parabolic Anderson. La posibilidad de perturbar ahora la ecuación del calor pero con otros procesos Gaussianos, tales como un movimiento browniano fraccionario, una sábana browniana, o incluso pensar en la posibilidad de perturbar la ecuación con procesos con saltos, procesos de Levy, y otros. Dada la naturaleza de estos últimos procesos, las técnicas del cálculo de Malliavin para distribuciones Poisson serían fundamentales para llevar a cabo el estudio.
- **Versiones funcionales.** Ya en [9], se puede vislumbrar una versión funcional del teorema que da origen a la cota (4.1). En [4], se presentan más resultados del tipo teoremas central del límite cuantitativos para los promedios espaciales de la solución a la ecuación del calor estocástica. En particular, se establecen versiones funcionales de dichos teoremas centrales del límite, además de un teorema central del límite para la solución *Hopf-Cole* de la ecuación *KPZ*. Por lo tanto, una dirección en la cual reflexionar sobre trabajo futuro sería llevar la convergencia de las densidades a una versión funcional.
- **Otros objetos que resuman información de la solución.** Se ha trabajado en este texto, y en todas las referencias dadas arriba, con los promedios espaciales de la solución a una SPDE. Aunque ciertamente los promedios espaciales resultan ser en cierta forma similares a una suma i.i.d. de variables independientes, y por lo tanto es esperable un comportamiento similar al teorema central de límite usual, otros objetos podrían comportarse de manera similar.

Como es de notar, el uso de las técnicas del cálculo de Malliavin junto con el método de Stein para deducir teoremas límite en el contexto de ecuaciones diferenciales parciales estocásticas es esencial, manteniéndose como una área bastante activa en la actualidad en el mundo de la probabilidad.

Apéndice A

Resultados técnicos

En este apéndice colocamos los tres resultados técnicos utilizados en la tercera parte de la prueba del resultado principal [39](#) de este texto. Específicamente, los lemas A.1, A.2 y A.3 de este apéndice corresponden respectivamente a los lemas A.2 parte (a), A.3 parte (a) y A.4 parte (a), del apéndice presente en el artículo de Şefika Kuzgun y David Nualart [\[11\]](#). Las pruebas del lema A.1 y A.3 se pueden encontrar en el artículo antes mencionado, mientras que la prueba del lema A.2 se puede encontrar en [\[9, proposición 3.1\]](#).

Lema A.1. Sea $\phi_{r,z,s,y}(t, x)$ de acuerdo al enunciado del teorema [41](#) y sea $(t, x) \in [0, \infty) \times \mathbb{R}$. Entonces para $0 < r < s < t$,

$$\int_{\mathbb{R}^2} \phi_{r,z,s,y}(t, x) dy dz \leq C_t \left(1 + \frac{1}{(s-r)^{1/4}} \right).$$

Lema A.2. Sea $\sigma_{R,t}$ según la definición [25](#). Entonces se cumple que

$$\lim_{R \rightarrow \infty} \frac{\sigma_{R,t}^2}{R} = 2 \int_0^t \xi(s) ds,$$

en donde $\xi(s) = \mathbb{E}[\sigma^2(u(s, y))]$, donde tal y como la notación indica, la función ξ es independiente del punto $y \in \mathbb{R}$ y además es acotada en intervalos compactos.

Lema A.3. Sean $t > 0$ y $\phi_{R,t}$ según se define en [\(4.14\)](#). Entonces existe un $R_0 \geq 1$, dependiente de t , tal que para cualesquiera $0 < s < t$ y $R \geq R_0$,

$$c_t \leq \int_{\mathbb{R}} \phi_{R,t}^2(s, y) dy \leq C_t,$$

en donde la cota inferior es válida siempre que $t/2 < s < t$.

Bibliografía

- [1] T. Bühler and D.A. Salamon, *Functional analysis*, Graduate Studies in Mathematics, American Mathematical Society, 2018.
- [2] María Emilia Caballero, Begoña Fernández, and David Nualart, *Estimation of densities and applications*, Journal of Theoretical Probability **11** (1998), no. 3, 831–851.
- [3] Le Chen and Robert Dalang, *Moments, intermittency and growth indices for the nonlinear fractional stochastic heat equation*, Stochastic Partial Differential Equations: Analysis and Computations **3** (2014).
- [4] Le Chen, Davar Khoshnevisan, David Nualart, and Fei Pu, *Central limit theorems for spatial averages of the stochastic heat equation via malliavin–stein’s method*, Stochastics and Partial Differential Equations: Analysis and Computations **11** (2023), no. 1, 122–176.
- [5] Masahisa Ebina, *Large deviations for a spatial average of stochastic heat and wave equations*, 2024.
- [6] L.C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, American Mathematical Society, 2010.
- [7] Julián Fernández Bonder, *Ecuaciones diferenciales parciales*, Cursos de grado, Departamento de Matemática, FCEyN, Universidad de Buenos Aires, 2015.
- [8] 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.
- [9] 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.
- [10] Davar Khoshnevisan, *A primer on stochastic partial differential equations*, pp. 1–38, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [11] Sefika Kuzgun and David Nualart, *Convergence of densities of spatial averages of stochastic heat equation*, Stochastic Processes and their Applications **151** (2022), 68–100.

- [12] ———, *Convergence of densities of spatial averages of the parabolic anderson model driven by colored noise*, *Stochastics* **96** (2024), no. 2, 968–984.
- [13] Jean-François Le Gall, *Brownian motion, martingales, and stochastic calculus*, Graduate Texts in Mathematics, Springer International Publishing, 2016.
- [14] Ivan Nourdin and Giovanni Peccati, *Normal approximations with malliavin calculus: From stein's method to universality*, Cambridge Tracts in Mathematics, Cambridge University Press, 2012.
- [15] David Nualart, *The malliavin calculus and related topics*, Probability and Its Applications, Springer Berlin Heidelberg, 2006.
- [16] David Nualart and Lluís Quer-Sardanyons, *Existence and smoothness of the density for spatially homogeneous spdes*, *Potential Analysis* **27** (2007), no. 3, 281–299.
- [17] Ciprian A Tudor and Jérémy Zurcher, *The spatial average of solutions to spdes is asymptotically independent of the solution*, 2024.
- [18] 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.
- [19] Wanying Zhang, Yong Zhang, and Jingyu Li, *Convergence of densities of spatial averages of the linear stochastic heat equation*, arXiv preprint arXiv:2308.03641 (2023).