



Sección de Matemáticas
Universidad de La Laguna

Raúl Melián Moreno

Introducción a las Ecuaciones Diferenciales Estocásticas y Métodos Numéricos.

Introduction to Stochastic Differential Equations
and numerical methods.

Trabajo de Fin de Grado
Grado en Matemáticas
La Laguna, Julio de 2019

Agradecimientos

Agradecer a mi tutor Domingo Hernández Abreu por su paciencia, profesionalidad y dedicación en este proyecto.

Gracias a mis padres por su gran apoyo incondicional para mí.

Gracias a mis amigos por ayudarme y estar siempre conmigo en esta etapa de mi vida.

Raúl Melián Moreno
La Laguna, 8 de julio de 2019

Resumen · Abstract

Resumen

En la presente memoria se pretende estudiar las ecuaciones diferenciales estocásticas. Inicialmente daremos una pequeña introducción a la teoría de probabilidad que es necesaria para su desarrollo. Seguidamente, definiremos el concepto de proceso de Wiener junto con sus propiedades. El objetivo será construir un proceso de Wiener que nos servirá para modelar componentes aleatorias en tales ecuaciones diferenciales. A continuación presentamos la definición y propiedades de la integral estocástica de Itô para posteriormente introducir la noción de solución de una ecuación diferencial estocástica, estableciendo la existencia y unicidad de dicha solución. Finalmente, veremos algunos ejemplos y aplicaciones de métodos numéricos tales como los método de Euler-Maruyama y Milsten.

Palabras clave: *Proceso de Wiener – Integral Estocástica de Itô – Ecuación diferencial Estocástica – Método de Euler-Maruyama – Método de Milstein.*

Abstract

In the present manuscript, we study stochastic differential equations. Initially, we will give a short introduction to the probability theory that is necessary for the development of these equations. Next, we will introduce the concept of Wiener process together with its properties. The objective will be to build a Wiener process that will allow us to model random components for such differential equations. Next, we will present the definition and properties of the Itô stochastic integral which will lead us to the concept of solution of a stochastic differential equation, establishing the existence and uniqueness of such solution. Finally, we will see some examples and applications of numerical methods such as the Euler-Maruyama and the Milsten methods.

Keywords: *Wiener process – Stochastic integral of Itô – Stochastic differential equation – Euler-Maruyama method – Milstein method.*

Contenido

Agradecimientos	III
Resumen/Abstract	V
1. Introducción y preliminares	1
1.1. Introducción	1
1.2. Preliminares sobre teoría de probabilidades	1
1.3. Variables aleatorias	3
1.4. Probabilidad condicionada e independencia	7
1.5. Esperanza condicionada y martingalas	9
2. Procesos de Wiener	11
2.1. Caminos aleatorios y la distribución normal	11
2.2. Procesos de Wiener: definición y propiedades	12
2.3. Construcción de un proceso de Wiener	13
2.4. Propiedades de las trayectorias	17
3. La integral estocástica de Itô	21
3.1. Introducción a la integración estocástica	21
3.2. Procesos estocásticos escalonados	23
3.3. La integral de Itô	26
3.4. Diferenciales estocásticas: reglas de la cadena y del producto de Itô ..	28
4. Ecuaciones diferenciales estocásticas	35
4.1. Definición y ejemplos	35
4.2. Existencia y unicidad de solución	39
4.3. Métodos numéricos elementales para EDEs	44
4.3.1. El método de Euler-Maruyama	44
4.3.2. El método de Milstein	47
4.3.3. Ilustración numérica	49
Bibliografía	51
Póster	53

Introducción y preliminares

1.1. Introducción

En este trabajo daremos una introducción a las ecuaciones diferenciales estocásticas. Son ecuaciones diferenciales que contienen una parte determinista y un término adicional aleatorio y se utilizan para modelar y estudiar diversos fenómenos que conllevan incertidumbre. A mediados del siglo 20, fueron desarrolladas, entre otros, por los matemáticos Itô, Stratonovich y Langevin.

Al poseer estas ecuaciones términos aleatorios, el cálculo clásico no es el adecuado para tratarlas por lo que es necesario la introducción de un cálculo estocástico que dé sentido a dichos modelos y a las soluciones de las ecuaciones resultantes. Por ello, el objetivo de este trabajo será desarrollar la teoría de ecuaciones diferenciales estocásticas en el sentido de Itô con el objetivo final de establecer la existencia y unicidad de solución para tales ecuaciones, bajo condiciones similares a las usuales en el caso ordinario clásico.

Iniciaremos la memoria resumiendo algunas ideas básicas de probabilidad y procesos estocásticos. En el capítulo 2, particularmente definiremos el concepto de proceso de Wiener y se mencionarán algunas de sus propiedades. En el capítulo 3, se pretende definir y estudiar propiedades de las integrales estocásticas con respecto al proceso de Wiener en el sentido de Itô, además de las reglas de la cadena y del producto de Itô. Finalmente, en el capítulo final, definiremos la noción de solución de una ecuación diferencial estocástica, veremos algunos ejemplos y presentamos métodos numéricos para dicha aproximar dicha solución

1.2. Preliminares sobre teoría de probabilidades

Dado un experimento aleatorio se denomina espacio muestral Ω al conjunto de todos los posibles resultados de dicho experimento. Un suceso será un subconjunto de Ω . En general no siempre es posible asignar una medida de probabilidad sobre todos los subconjuntos del espacio muestral (a los que englobamos en el conjunto de las partes $\mathcal{P}(\Omega)$). Esto nos conduce al concepto de σ -álgebra.

Definición 1.1. Sea Ω un conjunto no vacío. $\mathcal{U} \subseteq \mathcal{P}(\Omega)$ es una σ -álgebra si verifica que

- (i) $\Omega \in \mathcal{U}$,
- (ii) $A \in \mathcal{U} \Rightarrow A^c \in \mathcal{U}$, donde $A^c := \Omega \setminus A$,
- (iii) $\{A_i\}_{i \in \mathbb{N}} \subset \mathcal{U} \Rightarrow \bigcup_{i \in \mathbb{N}} A_i \in \mathcal{U}$.

Definición 1.2. Sea Ω un conjunto no vacío y $\mathcal{U} \subseteq \mathcal{P}(\Omega)$ una σ -álgebra. Una función de probabilidad en (Ω, \mathcal{U}) es una aplicación $P : \Omega \rightarrow \mathbb{R}$ verificando

- (i) $P(\Omega) = 1$,
(ii) $P(A) \geq 0$, $\forall A \in \mathcal{U}$,
(iii) $P\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$, si $\{A_i\}_{i \in \mathbb{N}} \subset \mathcal{U}$, $A_i \cap A_j = \emptyset$, $\forall i \neq j$.

(Ω, \mathcal{U}, P) se denomina espacio probabilístico. Si A es un conjunto tal que $A \in \mathcal{U}$ decimos que A es un suceso del experimento aleatorio, mientras que si $w \in \Omega$ decimos que w es un punto de muestra. Asimismo, diremos que un suceso A es casi seguro (c.s.) si $P(A) = 1$.

Ejemplo 1.3. Sean $\Omega = \{w_1, w_2, \dots, w_N\}$ un conjunto finito, números reales $0 \leq p_j \leq 1$ para $j = 1, \dots, N$, tales que $\sum p_j = 1$, y $\mathcal{U} = \mathcal{P}(\Omega)$. Entonces para un subconjunto $A = \{w_{j_1}, w_{j_2}, \dots, w_{j_m}\} \in \mathcal{U}$, con $1 \leq j_1 < j_2 < \dots < j_m \leq N$, definimos $P(A) := p_{j_1} + p_{j_2} + \dots + p_{j_m}$. Entonces (Ω, \mathcal{U}, P) es un espacio de probabilidad.

Ejemplo 1.4. La menor σ -álgebra que contiene a los subconjuntos abiertos de \mathbb{R}^n se denomina σ -álgebra de Borel, denotada como \mathcal{B} . Ahora, si f es una función no negativa e integrable tal que $\int_{\mathbb{R}^n} f dx = 1$, definimos

$$P(B) := \int_B f(x) dx, \quad \forall B \in \mathcal{B}.$$

$(\mathbb{R}^n, \mathcal{B}, P)$ es un espacio probabilístico y f es la densidad de la medida de probabilidad P .

Ejemplo 1.5. Fijado un punto $x_0 \in \mathbb{R}^n$ definimos

$$P(B) := \begin{cases} 1 & \text{si } x_0 \in B \\ 0 & \text{si } x_0 \notin B \end{cases}, \quad \forall B \in \mathcal{B}.$$

Entonces $(\mathbb{R}^n, \mathcal{B}, P)$ es un espacio probabilístico y P se denomina la medida de Dirac concentrada en el punto x_0 .

Definición 1.6. Se llama límite inferior de una sucesión de sucesos $\{A_n\}_{n \in \mathbb{N}}$, y lo denotamos por $\underline{\lim} A_n$, al suceso que ocurre cuando ocurren todos los sucesos A_n excepto quizás un número finito de ellos, esto es

$$\underline{\lim} A_n := \bigcup_{k=1}^{\infty} \left\{ \bigcap_{n=k}^{\infty} A_n \right\}.$$

Se llama límite superior de una sucesión de sucesos $\{A_n\}_{n \in \mathbb{N}}$, y lo denotamos como $\overline{\lim} A_n$, al suceso que ocurre cuando ocurren infinitos sucesos A_n , es decir,

$$\overline{\lim} A_n := \bigcap_{k=1}^{\infty} \left\{ \bigcup_{n=k}^{\infty} A_n \right\}.$$

Además es fácil comprobar que $\underline{\lim} A_n \subseteq \overline{\lim} A_n$. En particular, si $\underline{\lim} A_n = \overline{\lim} A_n$ entonces la sucesión $\{A_n\}$ de sucesos se dice convergente. Este suceso límite se denota como $\lim A_n$ -

Definición 1.7. Una sucesión de sucesos $\{A_n\}_{n \in \mathbb{N}}$ se dice monótona creciente (respectivamente decreciente) si $A_n \subseteq A_{n+1}$ (respectivamente $A_n \supseteq A_{n+1}$), para todo $n \in \mathbb{N}$.

Observación 1.8. No es difícil probar que toda sucesión monótona de sucesos $\{A_n\}_{n \in \mathbb{N}}$ es convergente y que

$$\lim A_n = \begin{cases} \bigcap_{n=1}^{\infty} A_n, & \text{si } \{A_n\}_{n \in \mathbb{N}} \text{ es decreciente,} \\ \bigcup_{n=1}^{\infty} A_n, & \text{si } \{A_n\}_{n \in \mathbb{N}} \text{ es creciente.} \end{cases}$$

Recopilamos a continuación algunas propiedades elementales de toda probabilidad.

Teorema 1.9. *Sea (Ω, \mathcal{U}, P) un espacio de probabilidad.*

- i) Si $A, B \in \mathcal{U}$ y $A \cap B = \emptyset$, entonces $P(A \cup B) = P(A) + P(B)$.
- ii) $P(A^c) = 1 - P(A)$.
- iii) $P(\emptyset) = 0$.
- iv) Si $A, B \in \mathcal{U}$ y $A \subseteq B$, entonces $P(B \setminus A) = P(B) - P(A)$.
- v) Si $A, B \in \mathcal{U}$ y $A \subseteq B$, entonces $P(B) \geq P(A)$.
- vi) $P(A) \leq 1$.
- vii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- viii) $P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$.
- ix) Si $\{A_n\}_{n \in \mathbb{N}}$ es convergente entonces $P(\lim A_n) = \lim_{n \rightarrow \infty} P(A_n)$.
- x) $P(\underline{\lim} A_n) \leq \lim_{n \rightarrow \infty} \inf P(A_n) \leq \lim_{n \rightarrow \infty} \sup P(A_n) \leq P(\overline{\lim} A_n)$.

Lema 1.10. (*Borel-Cantelli*). Si $\sum_{n=1}^{\infty} P(A_n) < \infty$, entonces $P(\overline{\lim} A_n) = 0$.

Demostración. Usando las propiedades v) y viii) del Teorema 1.9, para cada n tenemos que

$$P(\overline{\lim} A_n) \leq P\left(\bigcup_{m=n}^{\infty} A_m\right) \leq \sum_{m=n}^{\infty} P(A_m).$$

Como $\sum P(A_m) < \infty$, el término de la derecha tiende a cero cuando $n \rightarrow \infty$. □

1.3. Variables aleatorias

Definición 1.11. Se denomina variable aleatoria (v.a.) en (Ω, \mathcal{U}, P) a toda aplicación $X : \Omega \rightarrow \mathbb{R}^n$ verificando que $X^{-1}(B) \in \mathcal{U}$, para todo $B \in \mathcal{B}$. Equivalentemente decimos que X es \mathcal{U} -medible y denotamos $P(X^{-1}(B)) \equiv P(X \in B)$.

Entre las propiedades de las variables aleatorias destacamos las siguientes (ver [12, Teorema 1.8]).

Lema 1.12. (i) Toda aplicación constante, $X(w) = c$, para todo $w \in \Omega$, es una variable aleatoria.

(ii) Si X es una v.a. y $c \in \mathbb{R}$, entonces $X + c$ y $c \cdot X$ también lo son.

(iii) Si X e Y son v.a., entonces también lo son $X + Y$, $X - Y$, $X \cdot Y$. En el caso de que $Y(w) \neq 0$, entonces $\frac{X}{Y}$ es una v.a.

Ejemplo 1.13. Sea $A \in \mathcal{U}$. Entonces la función indicatriz de A

$$\mathcal{X}_A(w) := \begin{cases} 1 & \text{si } w \in A \\ 0 & \text{si } w \notin A, \end{cases}$$

es una variable aleatoria.

Ejemplo 1.14. Si $A_1, A_2, \dots, A_m \in \mathcal{U}$ son disjuntos dos a dos, con $\Omega = \cup_{i=1}^m A_i$, y $a_i \in \mathbb{R}$, $i = 1, \dots, m$, entonces

$$X = \sum_{i=1}^m a_i \mathcal{X}_{A_i} \quad (1.1)$$

es una variable aleatoria que denominamos escalonada.

Lema 1.15. Sea (Ω, \mathcal{U}, P) un espacio probabilístico, $X : \Omega \rightarrow \mathbb{R}^n$ una variable aleatoria y $\mathcal{U}(X) = \{X^{-1}(B)/B \in \mathcal{B}\} \subseteq \mathcal{U}$. Entonces $\mathcal{U}(X)$ es una σ -álgebra en Ω . \square

Observación 1.16. $\mathcal{U}(X)$ se denomina σ -álgebra generada por la v.a. X , y es la menor sub- σ -álgebra de \mathcal{U} respecto a la cual X es medible.

Definición 1.17. Un proceso estocástico en (Ω, \mathcal{U}, P) es una aplicación $X : [0, +\infty) \times \Omega \rightarrow \mathbb{R}^n$ tal que $X(t, \cdot)$ es una v.a. $\forall t \in [0, +\infty)$. Dado $w \in \Omega$, la aplicación $X(\cdot, w)$ se denomina trayectoria del proceso estocástico.

Definición 1.18. Sea $X = \sum_{i=1}^k a_i \mathcal{X}_{A_i}$ una variable aleatoria real escalonada, $X : \Omega \rightarrow \mathbb{R}$. Definimos la integral de X como $\int_{\Omega} X dP := \sum_{i=1}^k a_i P(A_i)$. Si X es una variable aleatoria real no negativa, definimos

$$\int_{\Omega} X dP := \sup_{\substack{Y \leq X, \\ Y \text{ escalonada}}} \int_{\Omega} Y dP.$$

Finalmente, si $X : \Omega \rightarrow \mathbb{R}$ variable aleatoria real arbitraria, se define

$$\int_{\Omega} X dP := \int_{\Omega} X^+ dP - \int_{\Omega} X^- dP$$

donde, $X^+ := \max(X, 0)$, $X^- := \max(-X, 0)$, de modo que $X = X^+ - X^-$.

Observación 1.19. En particular, $\int_{\Omega} 1 dP = P(\Omega) = 1$. Además, si $X : \Omega \rightarrow \mathbb{R}^n$ es una variable aleatoria n -dimensional con $X = (X^1, X^2, \dots, X^n)$, entonces escribimos $\int_{\Omega} X dP = (\int_{\Omega} X^1 dP, \dots, \int_{\Omega} X^n dP)$.

Definición 1.20. Sea $X : \Omega \rightarrow \mathbb{R}^n$ una variable aleatoria. Se llama esperanza y varianza de X a $E[X] := \int_{\Omega} X dP$, $V[X] := \int_{\Omega} (X - E[X])^2 dP$, respectivamente. En particular, $V[X] = E[|X|^2] - |E[X]|^2$, donde el cuadrado en la integral y $|\cdot|$ denotan el producto escalar y norma euclídea en \mathbb{R}^n , respectivamente.

Definición 1.21. Sea $0 < p < \infty$. Se define el espacio $L^p(\Omega)$ como

$$L^p(\Omega) := \{X \text{ v.a. } / E[|X|^p] < \infty\}.$$

Observación 1.22. Si $1 \leq p < \infty$ entonces la aplicación $\|X\|_{L^p(\Omega)} := E[|X|^p]^{\frac{1}{p}}$ es una norma en $L^p(\Omega)$. Además el espacio $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ es completo, esto es toda sucesión de Cauchy en dicho espacio es convergente (véase [12, Teorema 3.11]).

En particular, cuando $p = 2$, $\langle X, Y \rangle_{L^2(\Omega)} := E[X \cdot Y]$ define un producto escalar en $L^2(\Omega)$ y se verifica la desigualdad de Cauchy-Schwartz

$$|\langle X, Y \rangle_{L^2(\Omega)}| \leq \|X\|_{L^2(\Omega)} \cdot \|Y\|_{L^2(\Omega)}. \tag{1.2}$$

Lema 1.23. (*Desigualdad de Chebyshev*). Sea X una v.a. y $p \geq 0$. Entonces

$$P(|X| \geq \lambda) \leq \frac{1}{\lambda^p} E[|X|^p] \quad \text{para todo } \lambda > 0. \tag{1.3}$$

Demostración. $E[|X|^p] = \int_{\Omega} |X|^p dP \geq \int_{\{|X| \geq \lambda\}} |X|^p dP \geq \lambda^p P(|X| \geq \lambda)$. □

Definición 1.24. Sea $X : \Omega \rightarrow \mathbb{R}^n$ una variable aleatoria. Entonces la función de distribución de X es la función $F_X : \mathbb{R}^n \rightarrow [0, 1]$ definida por $F_X(x) := P(X \leq x)$ para todo $x \in \mathbb{R}^n$. Si $X_1, \dots, X_m : \Omega \rightarrow \mathbb{R}^n$ son variables aleatorias, entonces su función de distribución conjunta es $F_{X_1, \dots, X_m} : (\mathbb{R}^n)^m \rightarrow [0, 1]$ definida como $F_{X_1, \dots, X_m}(x_1, \dots, x_m) := P(X_1 \leq x_1, \dots, X_m \leq x_m)$ para todo $x_k \in \mathbb{R}^n, k = 1, \dots, m$.

Definición 1.25. Sea $X : \Omega \rightarrow \mathbb{R}^n$ una variable aleatoria y $F = F_X$ su función de distribución. Una función no negativa e integrable (Lebesgue) $f : \mathbb{R}^n \rightarrow \mathbb{R}$ se dice función densidad de X si

$$F(x) = F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(y_1, \dots, y_n) dy_n \dots dy_1, \quad \forall x \in \mathbb{R}^n.$$

En particular, X se dice absolutamente continua y

$$P(X \in B) = \int_B f(x) dx, \quad \text{para todo } B \in \mathcal{B}. \tag{1.4}$$

Ejemplo 1.26. La función de densidad

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|x-\mu|^2}{2\sigma^2}} \quad x \in \mathbb{R},$$

define una variable aleatoria normal de media μ y varianza σ^2 , que denotamos como $X \sim N(\mu, \sigma^2)$. En efecto,

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu \quad \text{y} \quad \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sigma^2.$$

Lema 1.27. Sea $X : \Omega \rightarrow \mathbb{R}^n$ una variable aleatoria con función de densidad $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Si $g : \mathbb{R}^n \rightarrow \mathbb{R}$ es integrable (Lebesgue). Entonces

$$E[g(X)] = \int_{\mathbb{R}^n} g(x) f(x) dx.$$

En particular, $E[X] = \int_{\mathbb{R}^n} x f(x) dx$ y $V[X] = \int_{\mathbb{R}^n} (x - E[X])^2 f(x) dx$.

Demostración. Suponemos en primer lugar que g es una función escalonada en \mathbb{R}^n , esto es, $g = \sum_{i=1}^m b_i \mathcal{X}_{B_i}$ con $B_i \in \mathcal{B}$. Entonces

$$E[g(X)] = \sum_{i=1}^m b_i \int_{\Omega} \mathcal{X}_{B_i}(X) dP = \sum_{i=1}^m b_i P(X \in B_i) = \sum_{i=1}^m b_i \int_{B_i} f(x) dx = \int_{\mathbb{R}^n} g(x) f(x) dx.$$

Finalmente, para una función g integrable arbitraria, tendremos como consecuencia de lo anterior que

$$E[g(X)] = \sup_{Y \leq g(X)} \int_{\Omega} Y dP = \sup_{\phi \leq g} \int_{\Omega} \phi(X) dP = \sup_{\phi \leq g} \int_{\mathbb{R}^n} \phi(X) f(x) dx = \int_{\mathbb{R}^n} g(X) f(x) dx,$$

donde Y y ϕ son v.a. y funciones escalonadas. \square

Definición 1.28. (*Convergencia de variables aleatorias*). Sea una sucesión de variables aleatorias $X_1, X_2, \dots, X_n, \dots$ en \mathbb{R} definidas en un mismo espacio probabilístico (Ω, \mathcal{U}, P) . Podemos definir varios tipos de convergencia:

(I) *Convergencia casi seguro (c.s.)*. $X_n \xrightarrow[n \rightarrow \infty]{c.s.} X$ si

$$P\left(\{w \in \Omega : \lim_{n \rightarrow \infty} |X_n(w) - X(w)| = 0\}\right) = 1. \quad (1.5)$$

(II) *Convergencia en media cuadrática*. $X_n \xrightarrow[n \rightarrow \infty]{L^2(\Omega)} X$ si

$$E[X_n^2] < \infty, n \geq 1, E[X^2] < \infty \quad y \quad \lim_{n \rightarrow \infty} E[(X_n - X)^2] = 0. \quad (1.6)$$

(III) *Convergencia en probabilidad*. $X_n \xrightarrow[n \rightarrow \infty]{P} X$ si

$$\lim_{n \rightarrow \infty} P(\{w \in \Omega : |X_n(w) - X(w)| \geq \varepsilon\}) = 0, \quad \text{para todo } \varepsilon > 0. \quad (1.7)$$

(IV) *Convergencia en distribución*. $X_n \xrightarrow[n \rightarrow \infty]{D} X$ si

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x) \quad \text{para todo punto de continuidad } x \text{ de } F_X. \quad (1.8)$$

Teorema 1.29. *Respecto a la convergencia de variables aleatorias se tiene que*

i) (I) \Rightarrow (III).

ii) (II) \Rightarrow (III).

iii) (III) \Rightarrow (I) para alguna subsucesión, esto es, existe una subsucesión $\{X_{k_j}\}_{j=1}^{\infty}$ tal que $X_{k_j} \xrightarrow[j \rightarrow \infty]{c.s.} X$.

iv) (III) \Rightarrow (IV).

Demostración. i) Sea $B := \{w \in \Omega / \lim_{n \rightarrow \infty} X_n(w) \neq X(w)\}$. Por hipótesis, $P(B) = 0$.

Sean $\varepsilon > 0$ y

$$A_n := \bigcup_{m \geq n} \{w : |X_m(w) - X(w)| > \varepsilon\}.$$

$\{A_n\}_{n=1}^{\infty}$ es una sucesión decreciente y converge al suceso $A := \bigcap_{n=1}^{\infty} A_n$.

Si $w \notin B$ entonces existe $N = N(w) \in \mathbb{N}$ tal que $|X_n(w) - X(w)| \leq \varepsilon, \forall n \geq N$.

Luego, $w \notin A_n, \forall n \geq N$ y por tanto $w \notin A$. Esto implica que $A \subseteq B$. Luego,

$$0 = P(A) = \lim_{n \rightarrow \infty} P(A_n).$$

- ii) Es inmediata a partir del Lema 1.23 con $p = 2$.
- iii) Para cada $j \geq 1$, sea $\varepsilon_j = \frac{1}{j}$. Por hipótesis, existe $k_j \in \mathbb{N}$ tal que

$$P\left(|X_{k_j} - X| \geq \frac{1}{j}\right) \leq \frac{1}{j^2}.$$

En particular podemos tomar $k_1 < k_2 < \dots$ ($k_j \rightarrow \infty, j \rightarrow \infty$). Sean $A_j = \{w \in \Omega : |X_{k_j}(w) - X(w)| \geq \frac{1}{j}\}$, de modo que $\sum_{j=1}^{\infty} P(A_j) \leq \sum_{j=1}^{\infty} \frac{1}{j^2} < \infty$. Luego, según el Lema

1.10 $P(\bigcap_{l=1}^{\infty} \bigcup_{j=l}^{\infty} A_j) = 0$, o bien $P(\bigcup_{l=1}^{\infty} \bigcap_{j=l}^{\infty} A_j^c) = 1$. Sea $B := \bigcup_{l=1}^{\infty} \bigcap_{j=l}^{\infty} A_j^c$. Entonces:

$$w \in B \Leftrightarrow \exists J(w) \geq 1 : |X_{k_j}(w) - X(w)| < \frac{1}{j}, \forall j \geq J \Leftrightarrow \lim_{j \rightarrow \infty} |X_{k_j}(w) - X(w)| = 0.$$

En definitiva, como $P(B) = 1$, $P(\{w \in \Omega : \lim_{j \rightarrow \infty} |X_{k_j}(w) - X(w)| = 0\}) = 1$.

- (iv) Es inmediato comprobar que para cualesquiera v.a. Y, Z y números reales x y $\epsilon > 0$:

$$\{w : Y(w) \leq x\} \subseteq \{w : Z(w) \leq x + \epsilon\} \cup \{w : |Y(w) - Z(w)| > \epsilon\}.$$

Usando esta propiedad sigue que

$$\begin{aligned} P(X_n \leq x) &\leq P(X \leq x + \epsilon) + P(|X_n - X| > \epsilon) \quad y \\ P(X \leq x - \epsilon) &\leq P(X_n \leq x) + P(|X_n - X| > \epsilon). \end{aligned}$$

Luego

$$P(X \leq x - \epsilon) - P(|X_n - X| > \epsilon) \leq P(X_n \leq x) \leq P(X \leq x + \epsilon) + P(|X_n - X| > \epsilon).$$

El resultado sigue tomando $n \rightarrow \infty, \epsilon \rightarrow 0^+$ y teniendo en cuenta que x es un punto de continuidad de F_X . □

1.4. Probabilidad condicionada e independencia

Definición 1.30. Sea (Ω, \mathcal{U}, P) un espacio probabilístico y dos sucesos $A, B \in \mathcal{U}$. La probabilidad del suceso A condicionada al suceso B es

$$P(A|B) := \frac{P(A \cap B)}{P(B)}, \quad \text{si } P(B) > 0.$$

Definición 1.31. Se dice que los sucesos A y B son independientes si

$$P(A \cap B) = P(A)P(B).$$

Además, los sucesos A_1, \dots, A_n, \dots son independientes si para cualesquiera $1 \leq k_1 < k_2 < \dots < k_m$ se tiene que

$$P(A_{k_1} \cap A_{k_2} \cap \dots \cap A_{k_m}) = P(A_{k_1})P(A_{k_2}) \cdots P(A_{k_m}).$$

Definición 1.32. Sean $\mathcal{U}_i \subseteq \mathcal{U}$ σ -álgebras, para $i \geq 1$. Se dice que $\{\mathcal{U}_i\}_{i=1}^{\infty}$ son independientes si para cualesquiera $1 \leq k_1 < k_2 < \dots < k_m$ y cualesquiera sucesos $A_{k_i} \in \mathcal{U}_{k_i}$ se tiene que

$$P(A_{k_1} \cap A_{k_2} \cap \dots \cap A_{k_m}) = P(A_{k_1})P(A_{k_2}) \cdots P(A_{k_m}).$$

Definición 1.33. Sean $X_i : \Omega \rightarrow \mathbb{R}^n$ variables aleatorias, $i \geq 1$. Decimos que $\{X_i\}_{i=1}^{\infty}$ son independientes si para cualesquier $k \geq 2$ y cualesquiera subconjuntos de Borel $B_1, \dots, B_k \subseteq \mathbb{R}^n$, se tiene que

$$P(X_1 \in B_1, \dots, X_k \in B_k) = P(X_1 \in B_1) \cdots P(X_k \in B_k).$$

Esto es equivalente a decir que las σ -álgebras $\{\mathcal{U}(X_i)\}_{i=1}^{\infty}$ son independientes.

Recogemos a continuación (sin demostración) algunos resultados clásicos para variables aleatorias independientes. Para detalles sobre las demostraciones correspondientes remitimos al lector a [1, p. 36-39].

Teorema 1.34. Sean X_1, \dots, X_{l+k} v.a. en \mathbb{R}^n independientes y $f : \mathbb{R}^{nl} \rightarrow \mathbb{R}$, $g : \mathbb{R}^{nk} \rightarrow \mathbb{R}$ funciones medibles. Entonces $Y = f(X_1, \dots, X_l)$ y $Z = g(X_{l+1}, \dots, X_{l+k})$ son independientes.

Teorema 1.35. Las variables aleatorias $X_1, \dots, X_m : \Omega \rightarrow \mathbb{R}^n$ son independientes si y solo si

$$F_{X_1, \dots, X_m}(x_1, \dots, x_m) = F_{X_1}(x_1) \cdots F_{X_m}(x_m), \quad \forall x_k \in \mathbb{R}^n, k = 1, \dots, m.$$

Teorema 1.36. Si X_1, \dots, X_m son variables aleatorias en \mathbb{R} e independientes con $E[|X_i|] < \infty$, $i = 1, \dots, m$, entonces

$$E[|X_1 \cdots X_m|] < \infty \quad y \quad E[X_1 \cdots X_m] = E[X_1] \cdots E[X_m].$$

Además, si $V[|X_i|] < \infty$, $i = 1, \dots, m$, entonces

$$V[X_1 + \cdots + X_m] = V[X_1] + \cdots + V[X_m].$$

Definición 1.37. Sea X una v.a. en \mathbb{R}^n . Se llama función característica de X a

$$\phi_X(\lambda) := E[e^{i\lambda \cdot X}], \quad \lambda \in \mathbb{R}^n. \quad (1.9)$$

Ejemplo 1.38. Si $X \sim N(\mu, \sigma^2)$, entonces

$$\phi_X(\lambda) = e^{i\mu\lambda - \frac{\lambda^2\sigma^2}{2}}, \quad \lambda \in \mathbb{R}. \quad (1.10)$$

El siguiente resultado permite afirmar que toda v.a. queda determinada por su función característica. Para su demostración véase [1, p. 170-180]

Lema 1.39. (i) Si X_1, \dots, X_m son v.a. independientes, entonces para cada $\lambda \in \mathbb{R}^n$

$$\phi_{X_1 + \dots + X_m}(\lambda) = \phi_{X_1}(\lambda) \cdots \phi_{X_m}(\lambda).$$

(ii) Si X es una v.a. en \mathbb{R} , entonces $\phi^{(k)}(0) = i^k E[X^k]$, $k \geq 0$.

(iii) Si X e Y son v.a. y $\phi_X(\lambda) = \phi_Y(\lambda)$ para todo λ , entonces $F_X(x) = F_Y(x)$ para todo x .

Ejemplo 1.40. Si $X \sim N(\mu_1, \sigma_1^2)$ e $Y \sim N(\mu_2, \sigma_2^2)$ son independientes, entonces $X + Y$ es $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. En efecto,

$$\phi_{X+Y}(\lambda) = \phi_X(\lambda)\phi_Y(\lambda) = e^{i\mu_1\lambda - \frac{\lambda^2\sigma_1^2}{2}} e^{i\mu_2\lambda - \frac{\lambda^2\sigma_2^2}{2}} = e^{i(\mu_1+\mu_2)\lambda - \frac{\lambda^2}{2}(\sigma_1^2+\sigma_2^2)}.$$

Definición 1.41. Las variables aleatorias X_1, \dots, X_n, \dots en \mathbb{R} se dicen idénticamente distribuidas si $F_{X_1}(x) = F_{X_2}(x) = \dots = F_{X_n}(x) = \dots$ para todo x .

El siguiente resultado clásico (véase [1, p. 167-170]) permite afirmar que la suma normalizada $\frac{(X_1 + \dots + X_n) - n\mu}{\sqrt{n}\sigma}$ de v.a. $X_1 \dots X_n$ en \mathbb{R} idénticamente distribuidas con media μ y varianza σ^2 finitas, converge en distribución a una v.a. $N(0, 1)$.

Teorema 1.42. (Teorema Central del Límite) Sean X_1, \dots, X_n, \dots variables aleatorias en \mathbb{R} independientes e idénticamente distribuidas con $E[X_i] = \mu$ y $V[X_i] = \sigma^2 > 0$, para $i \geq 1$ y sea $S_n := X_1 + \dots + X_n, \forall n \geq 1$. Entonces para todos $-\infty < a < b < \infty$

$$\lim_{n \rightarrow \infty} P\left(a \leq \frac{S_n - n\mu}{\sqrt{n}\sigma} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx. \tag{1.11}$$

1.5. Esperanza condicionada y martingalas

Finalizamos este capítulo inicial con una somera introducción a los conceptos de esperanza condicionada para variables aleatorias y de martingala. En particular, nos interesa considerar la desigualdad de Doob para martingalas (1.12), pues juega un papel fundamental en la existencia de solución para ecuaciones diferenciales estocásticas (que introduciremos en el capítulo 4).

Definición 1.43. Sea X una variable aleatoria y B un suceso dado tal que $P(B) > 0$. Se llama esperanza de X condicionada a B a

$$E[X|B] = \frac{1}{P(B)} \int_B X dP.$$

Nuestro objetivo a continuación es establecer una definición de esperanza condicionada $E[X|Y]$ siendo X, Y v.a. Veamos para ello el siguiente ejemplo introductorio.

Ejemplo 1.44. Sea X una variable aleatoria e Y otra variable aleatoria escalonada conocida definida como (1.1), donde $A_1, A_2, \dots, A_m \in \mathcal{U}$ y a_1, \dots, a_m son números reales. Definimos la esperanza de X condicionada a Y como

$$E[X|Y] := \begin{cases} \frac{1}{P(A_1)} \int_{A_1} X dP, & \text{si } w \in A_1 \\ \frac{1}{P(A_2)} \int_{A_2} X dP, & \text{si } w \in A_2 \\ \vdots \\ \frac{1}{P(A_m)} \int_{A_m} X dP, & \text{si } w \in A_m. \end{cases}$$

Observamos que

(i) $E[X|Y]$ es una variable aleatoria no constante. Además $E[X|Y](w)$ representa la mejor aproximación para $X(w)$ si se conoce $Y(w)$.

(ii) $E[X|Y]$ es $\mathcal{U}(Y)$ -medible.

(iii) $\int_A X dP = \int_A E[X|Y] dP$ para todo $A \in \mathcal{U}(Y)$. En efecto, poniendo $A = \bigcup_{j=1}^k A_{ij}, 0 \leq k \leq m$, con $\{A_{ij}\}_{j=1}^k$ disjuntos dos a dos, tenemos que $\int_A E[X|Y] dP = \sum_{j=1}^k \int_{A_{ij}} E[X|Y] dP = \sum_{j=1}^k \int_{A_{ij}} \left[\frac{1}{P(A_{ij})} \int_{A_{ij}} X dP \right] dP = \sum_{j=1}^k \int_{A_{ij}} X dP = \int_A X dP$.

Esta propiedad (iii) permite dar una definición general de esperanza condicionada.

Definición 1.45. Sean X e Y variables aleatorias definidas en el mismo espacio probabilístico (Ω, \mathcal{U}, P) . La esperanza de X condicionada a Y es cualquier variable aleatoria $Z \mathcal{U}(Y)$ -medible tal que

$$\int_A X dP = \int_A Z dP \quad \text{para todo } A \in \mathcal{U}(Y).$$

Definición 1.46. Sea (Ω, \mathcal{U}, P) un espacio probabilístico, \mathcal{V} una σ -álgebra $\mathcal{V} \subseteq \mathcal{U}$ y $X : \Omega \rightarrow \mathbb{R}^n$ una variable aleatoria integrable. Se llama esperanza de X condicionada a \mathcal{V} , $E[X|\mathcal{V}]$, a cualquier variable aleatoria en Ω tal que

- (i) $E[X|\mathcal{V}]$ es \mathcal{V} -medible,
- (ii) $\int_A X dP = \int_A E[X|\mathcal{V}] dP$, para todo $A \in \mathcal{V}$.

Observación 1.47. Las esperanzas condicionadas $E[X|Y]$ y $E[X|\mathcal{V}]$ de las Definiciones 1.45 y 1.46 son únicas salvo en conjuntos de probabilidad cero (ver [3, p. 29-33]).

Lema 1.48. i) Si a, b son constantes, entonces $E[aX+bY|\mathcal{V}] = aE[X|\mathcal{V}] + bE[Y|\mathcal{V}]$ c.s.
 ii) Si X es \mathcal{V} -medible, entonces $E[X|\mathcal{V}] = X$ c.s.
 iii) Si X es independiente de \mathcal{V} , entonces $E[X|\mathcal{V}] = E[X]$ c.s.

Finalizamos este capítulo con la definición del concepto de martingala y la propiedad fundamental dada en el Teorema 1.51 que establecemos sin demostración. La prueba de dicho resultado escapa de los objetivos de este trabajo (véase [3, Sec. 2.6]).

Definición 1.49. Sea $X : [0, \infty) \times \Omega \rightarrow \mathbb{R}^n$ un proceso estocástico se denomina historia del proceso estocástico hasta el tiempo t , y lo denotamos por $\mathcal{U}(t)$, a la σ -álgebra generada por las variables aleatorias $X(s)$ para $0 \leq s \leq t$, esto es

$$\mathcal{U}(t) := \mathcal{U}(X(s) | 0 \leq s \leq t).$$

Definición 1.50. Sea $X : [0, \infty) \times \Omega \rightarrow \mathbb{R}^n$ un proceso estocástico tal que $E[|X(t)|] < \infty$ para todo $t \geq 0$. X es una martingala si

$$X(s) = E[X(t) | \mathcal{U}(s)], \quad \text{casi seguro para todo } t \geq s \geq 0.$$

Teorema 1.51. (Desigualdad de Doob para martingalas). Sea X una martingala con trayectorias continuas casi seguro y $1 < p < \infty$. Entonces

$$E \left[\max_{0 \leq s \leq t} |X(s)|^p \right] \leq \left(\frac{p}{p-1} \right)^p E[|X(t)|^p]. \quad (1.12)$$

Procesos de Wiener

2.1. Caminos aleatorios y la distribución normal.

Sean $\Delta x, \Delta t > 0$ cantidades positivas arbitrarias, pero inicialmente fijas, y la malla de puntos $\{(m\Delta x, n\Delta t)/m \in \mathbb{Z}, n \in \mathbb{N} \cup \{0\}\}$ en $\mathbb{R} \times [0, \infty)$. Consideramos una partícula puntual inicialmente en la posición $x = 0$ en tiempo $t = 0$. En un tiempo cualquiera de la forma $t = n\Delta t$, la partícula se desplaza una cantidad Δx a la izquierda o derecha con igual probabilidad $1/2$. Sea $P(m\Delta x, n\Delta t)$ la probabilidad de que la partícula esté en la posición $x = m\Delta x$ en tiempo $t = n\Delta t$. Entonces:

$$P(m\Delta x, 0) = \begin{cases} 0, & m = 0, \\ 1, & m = 1, \end{cases} \quad y$$

$$P(m\Delta x, n\Delta t + \Delta t) = \frac{P(m\Delta x - \Delta x, n\Delta t) + P(m\Delta x + \Delta x, n\Delta t)}{2}.$$

Asumiendo que $\frac{(\Delta x)^2}{\Delta t} = D$ es constante, se tiene para $x = m\Delta x$ y $t = n\Delta t$ que:

$$\frac{P(x, t + \Delta t) - P(x, t)}{\Delta t} = \frac{D}{2} \frac{P(x - \Delta x, t) - 2P(x, t) + P(x + \Delta x, t)}{(\Delta x)^2}. \quad (2.1)$$

Ahora, para $(x, t) \in \mathbb{R} \times [0, \infty)$ cualquiera, interesa conocer la densidad de probabilidad de que la partícula esté en la posición x en tiempo t . Para ello, hacemos $\Delta x \rightarrow 0, \Delta t \rightarrow 0$ en (2.1) y asumimos, en un sentido que será justificado posteriormente mediante el Teorema Central del Límite, que $P(m\Delta x, n\Delta t) \rightarrow u(x, t)$, cuando $\Delta x, \Delta t \rightarrow 0$, con $\frac{(\Delta x)^2}{\Delta t} = D$ constante.

De (2.1), asumiendo además que $u \in C^2$ respecto de $t > 0$ y $u \in C^4$ respecto de $x \in \mathbb{R}$, sigue que:

$$u_t = \frac{D}{2} u_{xx}, \quad t > 0, \quad x \in \mathbb{R}, \quad (2.2)$$

con una densidad inicial de probabilidades $u(x, 0)$ tal que $u(x, 0) = 0, \forall x \neq 0$, y $\int_{-\infty}^{\infty} u(x, 0) dx = 1$ (esto es, $u(x, 0) = \delta_0(x)$ es la medida de Dirac o masa unitaria concentrada en $x = 0$). Luego $u(x, t) = \frac{1}{\sqrt{2Dt}} e^{-\frac{x^2}{2Dt}}$, $x \in \mathbb{R}, t > 0$, solución fundamental de la ecuación de difusión (2.2). Así pues, para cada $t > 0$, $u(x, t)$ es la densidad de probabilidad de una variable aleatoria $U \sim N(0, Dt)$.

Veamos a continuación cómo justificar en (2.1) el paso al límite de una variable discreta a una variable continua mediante el Teorema 1.42. Sean para $i \in \mathbb{N}, i \geq 1$, las variables

$$X_i = \begin{cases} 0, & \text{si la partícula se desplaza a izquierda en tiempo } (i-1)\Delta t, \\ 1, & \text{si la partícula se desplaza a derecha en tiempo } (i-1)\Delta t. \end{cases} \quad (2.3)$$

Entonces $P(X_i = 0) = \frac{1}{2} = p(X_i = 1)$, $E[X_i] = \frac{1}{2}$ y $V[X_i] = \frac{1}{4}$, $\forall i \geq 1$.

Sea $S_n := \sum_{i=1}^n X_i$ la variable que cuenta el número de desplazamientos a la derecha en los n primeros pasos. Sea $t = n\Delta t$ y U_n la variable aleatoria que indica la posición de la partícula en dicho tiempo. Se tiene que $U_n = S_n\Delta x + (n - S_n)(-\Delta x) = (2S_n - n)\Delta x$. Además, por independencia $E[S_n] = \frac{n}{2}$, $V[S_n] = \frac{n}{4}$, $E[U_n] = 0$ y $V[U_n] = 4\left(\frac{n}{4}\right)(\Delta x)^2 = Dt$, donde en la última igualdad hemos usado que $(\Delta x)^2(\Delta t) = D$ constante. Por tanto:

$$P(a \leq U_n \leq b) = P\left(\frac{a}{\sqrt{Dt}} \leq \frac{2S_n - n}{\sqrt{n}} \leq \frac{b}{\sqrt{Dt}}\right) = P\left(\frac{a}{\sqrt{Dt}} \leq \frac{S_n - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} \leq \frac{b}{\sqrt{Dt}}\right).$$

Por el Teorema 1.42:

$$\lim_{n \rightarrow \infty} P(a \leq U_n \leq b) = \frac{1}{\sqrt{2\pi}} \int_{\frac{a}{\sqrt{Dt}}}^{\frac{b}{\sqrt{Dt}}} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi Dt}} \int_a^b e^{-\frac{x^2}{2Dt}} dx. \quad (2.4)$$

Así pues, $U_n \xrightarrow[n \rightarrow \infty]{} N(0, Dt)$ en distribución, para cada $t > 0$.

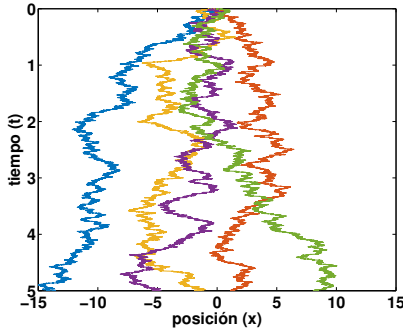


Figura 2.1. Cinco caminos aleatorios (2.3) con $D = 10$, $T = 5$, $\Delta t = \frac{T}{10^4}$ y $\Delta x = \sqrt{D\Delta t}$.

2.2. Procesos de Wiener: definición y propiedades

Teniendo en cuenta las consideraciones anteriores y con $D = 1$, definimos lo siguiente.

Definición 2.1. *Un proceso estocástico real W es un proceso de Wiener si:*

- (i) $W(0) = 0$ casi seguro.
- (ii) $W(t) - W(s)$ es $N(0, t - s)$, $\forall t \geq s \geq 0$.
- (iii) $W(t) - W(s)$ y $W(v) - W(u)$ son independientes para $0 \leq s < t < u < v$.

Observación 2.2. Si W es un proceso de Wiener, entonces para todo $t > 0$, ya que $W(t) \sim N(0, t)$

$$P(a \leq W(t) \leq b) = \frac{1}{\sqrt{2\pi t}} \int_a^b e^{-\frac{x^2}{2t}} dx.$$

Lema 2.3. *Sea W un proceso de Wiener. Entonces*

$$E[W(t)] = 0, \quad E[W^2(t)] = t, \quad \forall t \geq 0, \quad (2.5)$$

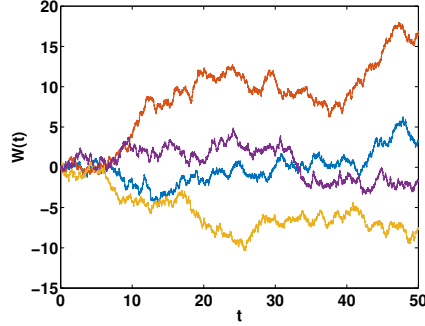


Figura 2.2. Cuatro trayectorias de un proceso de Wiener para $t \in [0, 50]$.

$$E[W(t)W(s)] = \min\{s, t\}, \quad \forall t, s \geq 0, \tag{2.6}$$

y

$$W \text{ es una martingala.} \tag{2.7}$$

Demostración. (2.5) es inmediato por la Observación (2.2). Veamos (2.6), para lo cual asumimos que $t \geq s \geq 0$. Entonces

$$\begin{aligned} E[W(t)W(s)] &= E[W^2(s)] + E[(W(t) - W(s))W(s)] \\ &= s + E[W(t) - W(s)]E[W(s)] = s = \min\{s, t\}, \end{aligned}$$

ya que $E[W(t) - W(s)] = 0$, $E[W(s)] = 0$ y $W(t) - W(s)$ es independiente de $W(s) - W(0)$.

Para ver (2.7), escribimos $\mathcal{W}(t) := \mathcal{U}(W(s)|0 \leq s \leq t)$, la σ -álgebra generada por W hasta tiempo $t = s$. Entonces, dado $t \geq s$, tendremos casi seguro usando la proposición 1.48 que

$$\begin{aligned} E[W(t)|\mathcal{W}(s)] &= E[W(t) - W(s)|\mathcal{W}(s)] + E[W(s)|\mathcal{W}(s)] \\ &= E[W(t) - W(s)] + W(s) \\ &= W(s). \end{aligned}$$

□

2.3. Construcción de un proceso de Wiener

Nuestro objetivo en esta sección es demostrar la existencia de un proceso de Wiener en el sentido de la Definición 2.1. Para ello consideraremos un desarrollo particular en serie de funciones ortonormales en el espacio

$$L^2(0, 1) := \left\{ f : [0, 1] \rightarrow \mathbb{R} / \int_0^1 |f(t)|^2 dt < \infty \right\}. \tag{2.8}$$

Es bien conocido que $L^2(0, 1)$ tiene estructura de espacio de Hilbert con el producto interior (véase [12, Cap. 3])

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt. \tag{2.9}$$

Definición 2.4. Un sistema $\{\psi(t)\}_{n \geq 0} \subset L^2(0, 1)$ se dice ortonormal respecto del producto (2.9) si

$$\langle \psi_n, \psi_m \rangle = \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases}$$

Además el sistema es completo si:

$$f \in L^2(0, 1) \text{ tal que } \langle f, \psi_n \rangle = 0, \forall n \geq 0 \Rightarrow f = 0 \text{ en } L^2(0, 1).$$

Definición 2.5. La familia de funciones de Haar $\{h_k(t)\}_{k=0}^\infty$ se define para $0 \leq t \leq 1$ como:

$$h_0(t) := 1, \quad 0 \leq t \leq 1, \quad h_1(t) := \begin{cases} 1, & 0 \leq t \leq \frac{1}{2}, \\ -1, & \frac{1}{2} < t \leq 1, \end{cases}$$

y si $2^n \leq k < 2^{n+1}$, con $n \geq 1$, entonces

$$h_k(t) := \begin{cases} 2^{\frac{n}{2}}, & \frac{k-2^n}{2^n} \leq t \leq \frac{k-2^n+\frac{1}{2}}{2^n}, \\ -2^{\frac{n}{2}}, & \frac{k-2^n+\frac{1}{2}}{2^n} < t \leq \frac{k-2^n+1}{2^n}, \\ 0, & \text{en otro caso.} \end{cases} \quad (2.10)$$

Lema 2.6. Las funciones de Haar $\{h_k(t)\}_{k=0}^\infty$ forman un sistema ortonormal completo de $(L^2(0, 1), \langle \cdot, \cdot \rangle)$.

Demostración. En primer lugar, veamos la ortonormalidad de las funciones de Haar $\forall k \geq 0$. Tenemos que $\int_0^1 h_k^2 dt = 2^n \left(\frac{1}{2^{n+1}} + \frac{1}{2^{n+1}} \right) = 1$ y además $\int_0^1 h_k dt = 0, \forall k \geq 1$.

Ahora, si $l > k$, entonces o bien $h_k(t)h_l(t) = 0$ para todo t , o h_k es constante en el soporte de h_l . En este último caso,

$$\int_0^1 h_l h_k dt = \int_{\text{sop}(h_l)} h_l h_k dt = h_k \int_{\text{sop}(h_l)} h_l dt = 0.$$

Veamos ahora la completitud del sistema. Suponemos que $f \in L^2(0, 1)$ es tal que $\int_0^1 f h_k dt = 0$ para todo $k = 0, 1, \dots$. Demostraremos que $f = 0$ en $L^2(0, 1)$. Para $k = 0$, tenemos $\int_0^1 f dt = 0$. Si $k = 1$, $\int_0^{\frac{1}{2}} f dt = \int_{\frac{1}{2}}^1 f dt$ y ambas son cero ya que $0 = \int_0^{\frac{1}{2}} f dt + \int_{\frac{1}{2}}^1 f dt = \int_0^1 f dt$. De manera similar deducimos que $\int_{\frac{k}{2^{n+1}}}^{\frac{k+1}{2^{n+1}}} f dt = 0$ para todo $0 \leq k < 2^{n+1}$, $n \geq 0$. Así, $\int_s^r f dt = 0$ para cualesquiera racionales diádicos $s, r \in [0, 1]$ y, en conclusión, para todos $0 \leq s \leq r \leq 1$ reales. Esto implica que $f = 0$ en casi todo punto de $[0, 1]$. \square

Definición 2.7. Dado $k \geq 0$ natural, se denomina k -ésima función de Schauder a

$$s_k(t) = \int_0^t h_k(s) ds, \quad 0 \leq t \leq 1, \quad (2.11)$$

siendo $\{h_k(t)\}_{k=0}^\infty$ las funciones de Haar (2.10). Para las funciones de Schauder observamos que

$$\max_{0 \leq t \leq 1} |s_0(t)| = 1 \quad y \quad \max_{0 \leq t \leq 1} |s_k(t)| = 2^{\frac{-n}{2}-1}, \text{ para } 2^n \leq k < 2^{n+1}, \quad n \geq 0. \quad (2.12)$$

Lema 2.8. *Las funciones de Schauder (2.11) verifican*

$$\sum_{k=0}^{\infty} s_k(t)s_k(s) = \min(t, s) \quad \text{para } 0 \leq t, s \leq 1.$$

Demostración. Definimos para cada $0 \leq t \leq 1$ las funciones

$$\phi_t(u) = \begin{cases} 1, & 0 \leq u \leq t, \\ 0, & t < u \leq 1. \end{cases} \quad (2.13)$$

Por el Lema 2.6, y teniendo en cuenta que todo sistema ortonormal completo es una base ([12, p. 85]), tendremos que

$$\phi_t = \sum_{k=0}^{\infty} a_k h_k, \quad \text{con } a_k = \langle \phi_t, h_k \rangle, \quad k \geq 0.$$

Como $a_k = \int_0^1 \phi_t(u)h_k(u)du = \int_0^t h_k(u)du = s_k(t)$, sigue de (2.13) y de la ortonormalidad de las funciones $\{h_k\}_{k=0}^{\infty}$ que

$$\min\{t, s\} = \int_0^1 \phi_t(u)\phi_s(u)du = \langle \phi_t, \phi_s \rangle = \sum_{k=0}^{\infty} s_k(t)s_k(s). \quad \square$$

Nuestra intención será definir un proceso de Wiener mediante

$$W(t, w) := \sum_{k=0}^{\infty} A_k(w)s_k(t), \quad 0 \leq t \leq 1, \quad (2.14)$$

donde los coeficientes $\{A_k(w)\}_{k=0}^{\infty}$ son variables aleatorias $N(0, 1)$ independientes definidas en un espacio de probabilidad (Ω, \mathcal{U}, P) . La expresión (2.14) requiere estudiar la convergencia uniforme de tal serie y para ello recurrimos a los dos siguientes lemas.

Lema 2.9. *Sea $\{a_k\}_{k=0}^{\infty}$ una sucesión de números naturales tal que*

$$|a_k| \leq Ck^{\delta}, \quad k = 1, 2, \dots,$$

con C constante y $0 \leq \delta < \frac{1}{2}$. Entonces la serie $\sum_{k=0}^{\infty} a_k s_k(t)$ converge uniformemente en $0 \leq t \leq 1$, siendo $\{s_k(t)\}_{k=0}^{\infty}$ la familia de funciones de Schauder (2.11).

Demostración. Fijemos $\varepsilon > 0$. Sabemos que para $2^n \leq k < 2^{n+1}$ las funciones s_k tienen soportes disjuntos. Sea

$$b_n := \max_{2^n \leq k < 2^{n+1}} |a_k| \leq C(2^{n+1})^{\delta}.$$

Entonces para $0 \leq t \leq 1$, (2.12) implica que

$$\sum_{k=2^m}^{\infty} |a_k||s_k(t)| \leq \sum_{n=m}^{\infty} b_n \max_{2^n \leq k < 2^{n+1}} |s_k(t)| \leq C \sum_{n=m}^{\infty} (2^{n+1})^{\delta} 2^{\frac{-n}{2}-1} < \varepsilon,$$

para m suficientemente grande, ya que $0 \leq \delta < \frac{1}{2}$ y la serie $\sum_{k=0}^{\infty} (2^{n+1})^{\delta} 2^{\frac{-n}{2}-1}$ es convergente. □

Lema 2.10. *Supongamos que $\{A_k\}_{k=1}^{\infty}$ son variables aleatorias $N(0, 1)$ independientes. Entonces, para casi todo w , $|A_k(w)| = \mathcal{O}(\sqrt{\log k})$, $k \rightarrow \infty$.*

Observación 2.11. Como $\sqrt{\log k} \leq Ck^{\frac{1}{4}}$, los números $\{A_k(w)\}_{k=1}^{\infty}$ satisfacen casi seguro la hipótesis del Lema 2.9.

Demostración. Para todo $x \geq 0$ y $k \geq 1$ tenemos que

$$P(|A_k| > x) = \frac{2}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{s^2}{2}} ds \leq \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{4}} \int_x^{\infty} e^{-\frac{s^2}{4}} ds \leq Ce^{-\frac{x^2}{4}},$$

con C constante. Tomando $x := 4\sqrt{\log k}$ tenemos que

$$P(|A_k| > 4\sqrt{\log k}) \leq Ce^{-4\log k} = C\frac{1}{k^4}.$$

Definiendo $S_k = \{w \in \Omega / |A_k(w)| > 4\sqrt{\log k}\}$, ya que $\sum_{k=1}^{\infty} \frac{1}{k^4} < \infty$, por el Lema 1.10

sabemos que $P(\bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} S_j) = 0$, esto es, $P(\bigcup_{k=1}^{\infty} \bigcap_{j=k}^{\infty} S_j^c) = 1$. Pero

$$w \in \bigcup_{k=1}^{\infty} \bigcap_{j=k}^{\infty} S_j^c \Leftrightarrow \exists K = K(w), K \in \mathbb{N}, \text{ tal que } |A_j(w)| \leq 4\sqrt{\log j}, \quad \forall j \geq K. \quad \square$$

Teorema 2.12. (*Construcción del proceso de Wiener*). Sea $\{A_k\}_{k=1}^{\infty}$ una sucesión de variables aleatorias $N(0, 1)$ independientes definidas en el mismo espacio de probabilidad (Ω, \mathcal{U}, P) . Entonces la serie

$$W(t, w) := \sum_{k=0}^{\infty} A_k(w) s_k(t), \quad 0 \leq t \leq 1, \quad w \in \Omega, \quad (2.15)$$

converge uniformemente en t para casi todo w . Además,

(i) W es un proceso de Wiener en $0 \leq t \leq 1$.

(ii) Para casi todo w , la trayectoria $t \mapsto W(t, w)$ es continua.

Demostración. La convergencia uniforme para (2.15) es consecuencia de los Lemas 2.9 y 2.10. Esto implica (ii).

Probemos que W es un proceso de Wiener. Observamos que $W(0) = 0$ casi seguro. Veamos que $W(t) - W(s)$ es $N(0, t - s)$ para todo $0 \leq s \leq t \leq 1$. Tenemos que

$$E[e^{i\lambda(W(t) - W(s))}] = E[e^{i\lambda \sum_{k=0}^{\infty} A_k(s_k(t) - s_k(s))}].$$

Teniendo en cuenta que $\{A_k\}_{k=0}^{\infty}$ son $N(0, 1)$ independientes y la expresión de la función característica de una variable aleatoria normal (1.10):

$$\begin{aligned} E[e^{i\lambda(W(t) - W(s))}] &= \prod_{k=0}^{\infty} E[e^{i\lambda A_k(s_k(t) - s_k(s))}] = \prod_{k=0}^{\infty} e^{-\frac{\lambda^2}{2}(s_k(t) - s_k(s))^2} \\ &= e^{-\frac{\lambda^2}{2} \sum_{k=0}^{\infty} (s_k(t) - s_k(s))^2} = e^{-\frac{\lambda^2}{2}(t-s)} \end{aligned}$$

donde en la última igualdad hemos usado el Lema 2.8. En virtud del Lema 1.39[(iii)] $W(t) - W(s)$ es $N(0, t - s)$. Finalmente veamos que $W(t) - W(s)$ y $W(v) - W(u)$ son independientes para $0 \leq s < t \leq u < v$. En primer lugar, un desarrollo análogo al anterior muestra que

$$E[e^{i(\lambda(W(t)-W(s))+\mu(W(v)-W(u)))}] = E[e^{i\lambda(W(t)-W(s))}]E[e^{i\mu(W(v)-W(u))}].$$

Teniendo en cuenta la unicidad de las funciones características, Lema 1.39[(iii)], esto implica para la función de distribución conjunta que:

$$F_{W(t)-W(s), W(v)-W(u)} = F_{W(t)-W(s)}F_{W(v)-W(u)},$$

y esto equivale por el Teorema 1.35 a decir que $W(t) - W(s)$ y $W(v) - W(u)$ son independientes. \square

El teorema previo se puede generalizar para construir un proceso de Wiener definido para todo $t \geq 0$.

Teorema 2.13. *Sea (Ω, \mathcal{U}, P) un espacio de probabilidad y una sucesión $\{A_k\}_{k \in \mathbb{N}}$ de variables aleatorias $N(0, 1)$ independientes. Entonces existe un proceso de Wiener unidimensional W definido para $w \in \Omega$ y $t \geq 0$.*

Aunque no detallamos la demostración, ésta se basa en tomar una familia numerable de sucesiones de variables aleatorias $\{A_k^{(n)}\}_{k=0}^\infty N(0, 1)$ independientes, para cada $n \geq 1$, y para cada una de ellas construir procesos de Wiener independientes $W^n(t)$, para $0 \leq t \leq 1$. Definiendo

$$\begin{cases} W(0) = 0, \\ W(t) = W(n-1) + W^n(t - (n-1)), \quad \text{para } n-1 < t \leq n, \end{cases}$$

se obtiene que W es un proceso de Wiener definido en $t \geq 0$.

2.4. Propiedades de las trayectorias

En esta sección demostraremos que para casi todo $w \in \Omega$ la trayectoria $t \rightarrow W(t, w)$ es uniformemente Hölder continua para cada exponente $0 < \gamma < \frac{1}{2}$, pero no lo es para $\gamma > \frac{1}{2}$. En particular $t \rightarrow W(t, w)$ no es diferenciable en ningún punto casi seguro y es de variación infinita en todo subintervalo cerrado $[a, b] \subseteq \mathbb{R}^+ \cup \{0\}$.

La herramienta fundamental que usaremos para probar la continuidad Hölder de las trayectorias de un proceso de Wiener es el Teorema de Kolmogorov que enunciaremos a continuación, y cuya prueba posponemos al final de esta sección.

Definición 2.14. *Sea $0 < \gamma \leq 1$. Una función $f : [0, T] \rightarrow \mathbb{R}$ se dice uniformemente Hölder continua con exponente $\gamma > 0$ si existe una constante K tal que:*

$$|f(t) - f(s)| \leq K|t - s|^\gamma, \quad \forall s, t \in [0, T]. \tag{2.16}$$

Decimos que f es Hölder continua con exponente $\gamma > 0$ en un punto s si existe una constante K tal que:

$$|f(t) - f(s)| \leq K|t - s|^\gamma, \quad \forall t \in [0, T].$$

Teorema 2.15. *(Kolmogorov) Sea X un proceso estocástico con trayectorias continuas casi seguro tal que:*

$$E[|X(t) - X(s)|^\beta] \leq C|t - s|^{1+\alpha}, \quad \forall t, s \geq 0, \tag{2.17}$$

con $\beta, \alpha > 0, C \geq 0$ constantes. Entonces para cada $T > 0$, para casi todo w , y para $0 < \gamma < \frac{\alpha}{\beta}$ existe una constante $K = K(w, \gamma, T)$ tal que

$$|X(t, w) - X(s, w)| \leq K|t - s|^\gamma, \quad \forall 0 \leq s, t \leq T. \tag{2.18}$$

Esto es, para casi todo w , la trayectoria $t \rightarrow X(t, w)$ es uniformemente Hölder continua en $[0, T]$, con exponente $0 < \gamma < \frac{\alpha}{\beta}$.

Observación 2.16. Decimos que un proceso estocástico \hat{X} es una versión de X si

$$P(X(t) = \hat{X}(t)) = 1, \quad \text{para todo } t \geq 0. \quad (2.19)$$

El Teorema de Kolmogorov se puede generalizar omitiendo la hipótesis de que el proceso X tenga trayectorias continuas casi seguro. Así, resulta que si X verifica (2.17), entonces tiene una versión \hat{X} (2.19) que, casi seguro, tiene trayectorias uniformemente Hölder continuas con exponente $0 < \gamma < \frac{\alpha}{\beta}$.

Teorema 2.17. *Para casi todo w y cualquier $T > 0$, la trayectoria $t \mapsto W(t, w)$ es uniformemente Hölder continua en $[0, T]$ para $0 < \gamma < \frac{1}{2}$.*

Demostración. Puesto que $W(t) - W(s) \sim N(0, t - s)$, tenemos para $m \geq 1$ y $0 \leq s < t \leq 1$ que $E[|W(t) - W(s)|^{2m}] = \frac{1}{\sqrt{2\pi r}} \int_{\mathbb{R}} |x|^{2m} e^{-\frac{|x|^2}{2r}} dx$, con $r := t - s > 0$. Haciendo el cambio $y = \frac{x}{\sqrt{r}}$ si que

$$E[|W(t) - W(s)|^{2m}] = \frac{1}{\sqrt{2\pi}} r^m \int_{\mathbb{R}} |y|^{2m} e^{-\frac{|y|^2}{2}} dy = Cr^m = C|t - s|^m.$$

Para $\beta = 2m$ y $\alpha = m - 1$, con $m \geq 1$, $m \in \mathbb{N}$, utilizando el Teorema 2.15 tenemos que el proceso W tiene trayectorias uniformemente Hölder continuas en $[0, T]$ para casi todo w con exponente $0 < \gamma < \frac{\alpha}{\beta} = \frac{1}{2} - \frac{1}{2m}$. \square

Veremos ahora que las trayectorias de un proceso de Wiener, con probabilidad uno, no son Hölder continuas en ningún punto para exponentes $\gamma > \frac{1}{2}$, y por lo tanto no son diferenciables en ningún punto.

Teorema 2.18. *Para cada $\frac{1}{2} < \gamma \leq 1$ y para casi todo w , la trayectoria $t \mapsto W(t, w)$ no es Hölder continua con exponente γ en ningún punto. En particular, para casi todo w , la trayectoria $t \mapsto W(t, w)$ no es diferenciable en ningún punto y tiene variación infinita en cualquier intervalo cerrado.*

Observación 2.19. Dada $f(x)$ definida en $[a, b]$ se define su variación en $[a, b]$ como

$$V_a^b(f) = \sup_{\mathcal{P}} \sum_{i=0}^{|\mathcal{P}|-1} |f(x_{i+1}) - f(x_i)|,$$

donde el supremo se toma sobre todas las particiones \mathcal{P} de $[a, b]$

Demostración. Basta considerar W proceso de Wiener definido en $t \in [0, 1]$. Sea $\gamma > \frac{1}{2}$ y N natural fijo tal que $N(\gamma - \frac{1}{2}) > 1$. Observamos que si $t \mapsto W(t, w)$ es Hölder continua en un punto $0 \leq s < 1$ con exponente γ , entonces existe una constante $K = K(s, w)$ tal que

$$|W(t, w) - W(s, w)| \leq K|t - s|^\gamma, \quad t \in [0, 1].$$

Así, para $n \gg 1$, dado $i = [ns] + 1$, con $i \leq n$, y $j = i, i + 1, \dots, i + N - 1$ tenemos que

$$\begin{aligned} \left| W\left(\frac{j}{n}, w\right) - W\left(\frac{j+1}{n}, w\right) \right| &\leq \left| W(s, w) - W\left(\frac{j}{n}, w\right) \right| + \left| W(s, w) - W\left(\frac{j+1}{n}, w\right) \right| \\ &\leq K \left(\left| s - \frac{j}{n} \right|^\gamma + \left| s - \frac{j+1}{n} \right|^\gamma \right) \leq \frac{M}{n^\gamma}, \end{aligned}$$

donde $M = K(N^\gamma + (N + 1)^\gamma)$, pues $j - ns \leq N$, $i \leq j \leq i + N - 1$. Por lo tanto,

$$w \in A_{M,n}^i := \left\{ w \in \Omega / \left| W\left(\frac{j}{n}, w\right) - W\left(\frac{j+1}{n}, w\right) \right| \leq \frac{M}{n^\gamma}, \quad \text{para } j = i, \dots, i+N-1 \right\},$$

para algún $1 \leq i \leq n$, para algún $M \geq 1$, y para todo n suficientemente grande. Por tanto, si $w \in \Omega$ es tal que la trayectoria $t \mapsto W(t, w)$ es Hölder continua en un punto $s \in [0, 1]$ con exponente $\gamma > \frac{1}{2}$ entonces

$$w \in A := \bigcup_{M=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \bigcup_{i=1}^n A_{M,n}^i.$$

Veremos ahora que A tiene probabilidad 0, lo cual concluye la primera parte del enunciado. En efecto, para todo k y M fijos tenemos por el Teorema 1.9 (x) que

$$\begin{aligned} P\left(\bigcap_{n=k}^{\infty} \bigcup_{i=1}^n A_{M,n}^i\right) &\leq \liminf_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n A_{M,n}^i\right) \leq \liminf_{n \rightarrow \infty} \sum_{i=1}^n P(A_{M,n}^i) \\ &\leq \liminf_{n \rightarrow \infty} n \left(P\left(\left|W\left(\frac{1}{n}\right)\right| \leq \frac{M}{n^\gamma}\right) \right)^N, \end{aligned}$$

ya que las variables aleatorias $W(\frac{j+1}{n}) - W(\frac{j}{n})$ son $N(0, \frac{1}{n})$ independientes. Ahora

$$P\left(\left|W\left(\frac{1}{n}\right)\right| \leq \frac{M}{n^\gamma}\right) = \sqrt{\frac{n}{2\pi}} \int_{-Mn^{-\gamma}}^{Mn^{-\gamma}} e^{-\frac{nx^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-Mn^{\frac{1}{2}-\gamma}}^{Mn^{\frac{1}{2}-\gamma}} e^{-\frac{y^2}{2}} dy \leq Cn^{\frac{1}{2}-\gamma},$$

con $C = \frac{2M}{\sqrt{2\pi}}$. Deducimos que $P\left(\bigcap_{n=k}^{\infty} \bigcup_{i=1}^n A_{M,n}^i\right) \leq \liminf_{n \rightarrow \infty} nC^N \left(n^{\frac{1}{2}-\gamma}\right)^N = 0$, ya que

$N(\gamma - \frac{1}{2}) > 1$. De aquí se deduce entonces inmediatamente que $P(A) = 0$. Por lo tanto, con probabilidad 1, la trayectoria $t \mapsto W(t, w)$ no es Hölder continua de exponente γ para ningún $s \in [0, 1]$.

Finalmente si $t \mapsto W(t, w)$ fuera diferenciable en s , entonces sería Hölder continua en s con exponente $\gamma = 1$, pero esto no se verifica casi seguro. Por otro lado, si $t \mapsto W(t, w)$ fuera de variación finita en algún subintervalo, sería casi seguro diferenciable en casi todo punto (véase [9, Sec. 32]). \square

Demostración. (Teorema 2.15). Fijemos $0 < \gamma < \frac{\alpha}{\beta}$ y pongamos $T = 1$. Vamos a probar en primer lugar que para casi todo $w \in \Omega$:

$$\left| X\left(\frac{i+1}{2^n}, w\right) - X\left(\frac{i}{2^n}, w\right) \right| \leq K \frac{1}{2^{n\gamma}}, \quad \text{para } 0 \leq i \leq 2^n - 1, \text{ con } n \geq 0, \quad (2.20)$$

donde $K = K(w)$ es una constante suficientemente grande. Definimos

$$A_n := \left\{ w \in \Omega / \left| X\left(\frac{i+1}{2^n}, w\right) - X\left(\frac{i}{2^n}, w\right) \right| > \frac{1}{2^{n\gamma}}, \text{ para algún natural } i, 0 \leq i < 2^n \right\}.$$

Entonces, por la Desigualdad de Chebyshev (1.3) y la hipótesis

$$\begin{aligned} P(A_n) &\leq \sum_{i=0}^{2^n-1} P\left(\left|X\left(\frac{i+1}{2^n}\right) - X\left(\frac{i}{2^n}\right)\right| > \frac{1}{2^{n\gamma}}\right) \\ &\leq \sum_{i=0}^{2^n-1} E\left[\left|X\left(\frac{i+1}{2^n}\right) - X\left(\frac{i}{2^n}\right)\right|^\beta\right] \left(\frac{1}{2^{n\gamma}}\right)^{-\beta} \leq C \sum_{i=0}^{2^n-1} \left(\frac{1}{2^n}\right)^{1+\alpha} \left(\frac{1}{2^{n\gamma}}\right)^{-\beta} \\ &= C2^{n(-\alpha+\gamma\beta)}. \end{aligned}$$

Como $-\alpha + \gamma\beta < 0$ deducimos que $\sum_{n=1}^{\infty} P(A_n) < \infty$ y el Lema 1.10 implica que $P(\overline{\lim} A_n) = 0$. Entonces para casi todo w existe $m = m(w) \in \mathbb{N}$ tal que para $n \geq m$

$$\left| X\left(\frac{i+1}{2^n}, w\right) - X\left(\frac{i}{2^n}, w\right) \right| \leq \frac{1}{2^{n\gamma}} \text{ para todo } i, 0 \leq i \leq 2^n - 1. \quad (2.21)$$

A partir de (2.21) se deduce (2.20) para todo $n \geq 0$ tomando $K = K(w)$ suficientemente grande.

Probaremos finalmente que (2.20) implica la continuidad Hölder uniforme en $[0, 1]$ con exponente γ para la trayectoria $t \mapsto W(t, w)$. Fijemos entonces $w \in \Omega$ para el que se cumpla (2.20). Tomemos $0 < t_1 < t_2 < 1$ racionales diádicos ($t_i = \frac{a_i}{2^{b_i}}$, con $\text{mcd}(a_i, 2^{b_i}) = 1$, $i = 1, 2$). Entonces $t := t_2 - t_1 \in (0, 1)$ y existe un único $n \in \mathbb{N}$, $n \geq 1$, tal que $t \in [\frac{1}{2^n}, \frac{1}{2^{n-1}})$. Dado este n , existen únicos i, j naturales tales que $t_1 \in (\frac{i-1}{2^n}, \frac{i}{2^n}]$, $t_2 \in [\frac{j}{2^n}, \frac{j+1}{2^n})$, con $i-1 \leq j$, $i \in \{1, \dots, 2^n\}$, $j \in \{0, \dots, 2^n - 1\}$.

Entonces $t = t_2 - t_1 \in [\frac{j-i}{2^n}, \frac{j-i+2}{2^n})$. Como $j-i \geq -1$ y $t \in [\frac{1}{2^n}, \frac{2}{2^n})$, sólo puede ser $j-i \in \{0, 1\}$. En cualquier caso, por (2.20):

$$\left| X\left(\frac{i}{2^n}, w\right) - X\left(\frac{j}{2^n}, w\right) \right| \leq K \left| \frac{i-j}{2^n} \right|^\gamma \leq K t^\gamma.$$

Además, teniendo en cuenta que t_1, t_2 son racionales diádicos y usando la expresión binaria de los números reales en $(0, 1)$, podemos escribir:

$$t_1 = \frac{i}{2^n} - \sum_{l=1}^{k_1} \frac{\alpha_l}{2^{n+l}}, \quad t_2 = \frac{j}{2^n} + \sum_{l=1}^{k_2} \frac{\beta_l}{2^{n+l}}, \quad \text{con } \alpha_l, \beta_l \in \{0, 1\},$$

y $k_1 = k_1(t_1)$, $k_2 = k_2(t_2) \in \mathbb{N}$. Como por (2.20)

$$\left| X\left(\frac{i}{2^n} - \sum_{l=1}^r \frac{\alpha_l}{2^{n+l}}, w\right) - X\left(\frac{i}{2^n} - \sum_{l=1}^{r-1} \frac{\alpha_l}{2^{n+l}}, w\right) \right| \leq K \left(\frac{1}{2^{n+r}}\right)^\gamma,$$

en virtud de la desigualdad triangular sigue que

$$|X(t_1, w) - X(\frac{i}{2^n}, w)| \leq \sum_{r=1}^{k_1} K \left(\frac{1}{2^{n+r}}\right)^\gamma \leq \frac{K}{2^{n\gamma}} \sum_{r=1}^{\infty} \left(\frac{1}{2^\gamma}\right)^r = \frac{C}{2^{n\gamma}},$$

con $C = K \frac{1}{2^{\gamma-1}} = C(\gamma, w)$. En particular, $|X(t_1, w) - X(\frac{i}{2^n}, w)| \leq C t^\gamma$.

Análogamente podemos comprobar que $|X(t_2, w) - X(\frac{j}{2^n}, w)| \leq C t^\gamma$. En definitiva, existe $C = C(\gamma, w)$ tal que para cualesquiera racionales diádicos $0 < t_1 < t_2 < 1$ se tiene que $|X(t_2, w) - X(t_1, w)| \leq C t^\gamma$, $t = t_2 - t_1$. Esta propiedad se extiende a cualesquiera números reales $0 \leq t_1 < t_2 \leq 1$ teniendo en cuenta que, por hipótesis, el proceso estocástico X tiene trayectorias continuas casi seguro. \square

La integral estocástica de Itô

3.1. Introducción a la integración estocástica

Sea W un proceso de Wiener unidimensional. En este capítulo pretendemos definir y estudiar propiedades de la integral estocástica $\int_0^T G dW$ en el sentido de Itô donde G pertenecerá a una clase adecuada de procesos estocásticos. Comenzamos en primer lugar motivando la integral de Itô del propio proceso de Wiener, esto es, $\int_0^T W dW$. El procedimiento a seguir se basa en definir sumas de Riemann para dicha integral y estudiar el paso al límite.

Definición 3.1. Sea \mathcal{P} una partición del intervalo $[0, T]$, $\mathcal{P} = \{0 = t_0 < t_1 < \dots < t_m = T\}$, y $\tau_k \in [t_k, t_{k+1}]$, $0 \leq k \leq m-1$. Se llama suma de Riemann de $\int_0^T W dW$ a:

$$R = \sum_{k=0}^{m-1} W(\tau_k)(W(t_{k+1}) - W(t_k)).$$

Lema 3.2. (Variación cuadrática). Sea $[a, b]$ un intervalo en $[0, \infty)$ y

$$\mathcal{P}^n := \{a = t_0^n < t_1^n < \dots < t_{m_n}^n = b\} \quad (3.1)$$

con $|\mathcal{P}^n| \rightarrow 0$, $n \rightarrow \infty$, particiones de $[a, b]$, siendo $|\mathcal{P}^n| = \max_{0 \leq j \leq m_n-1} t_{j+1}^n - t_j^n$. Entonces

$$\sum_{k=0}^{m_n-1} (W(t_{k+1}^n) - W(t_k^n))^2 \xrightarrow[n \rightarrow \infty]{L^2(\Omega)} b - a. \quad (3.2)$$

Observación 3.3. Observar que convergencia en $L^2(\Omega)$ implica convergencia en probabilidad, y ésta implica convergencia casi seguro al menos a través de subsucesiones (Teorema 1.29).

Demostración. Llamemos $Q_n := \sum_{k=0}^{m_n-1} (W(t_{k+1}^n) - W(t_k^n))^2$. Entonces

$$Q_n - (b - a) = \sum_{k=0}^{m_n-1} ((W(t_{k+1}^n) - W(t_k^n))^2 - (t_{k+1}^n - t_k^n)).$$

Por lo tanto, $E[(Q_n - (b - a))^2]$ es

$$\sum_{k=0}^{m_n-1} \sum_{j=0}^{m_n-1} E[((W(t_{k+1}^n) - W(t_k^n))^2 - (t_{k+1}^n - t_k^n)) \cdot ((W(t_{j+1}^n) - W(t_j^n))^2 - (t_{j+1}^n - t_j^n))].$$

Para $k \neq j$, el término correspondiente de la suma doble es nulo debido a que $W(t) - W(s) \sim N(0, t - s)$ para todo $t \geq s \geq 0$ y a la independencia de los incrementos del proceso de Wiener (Definición 2.1). Luego

$$E[(Q_n - (b - a))^2] = \sum_{k=0}^{m_n-1} E[(Y_k^2 - 1)^2](t_{k+1}^n - t_k^n)^2,$$

donde $Y_k = Y_k^n := \frac{W(t_{k+1}^n) - W(t_k^n)}{\sqrt{t_{k+1}^n - t_k^n}}$ es $N(0, 1)$. Así tendremos que

$$E[(Q_n - (b - a))^2] = C \sum_{k=0}^{m_n-1} (t_{k+1}^n - t_k^n)^2 \leq C|\mathcal{P}^n|(b - a) \rightarrow 0, \quad n \rightarrow \infty,$$

donde $C = E[(N(0, 1)^2 - 1)^2] < \infty$. \square

Observación 3.4. Este lema nos dice que podemos interpretar $dW \approx (dt)^{\frac{1}{2}}$ en sentido $L^2(\Omega)$.

Observación 3.5. Análogamente, el mismo desarrollo del lema previo nos dice que si $\tau_k^n = (1 - \lambda)t_k + \lambda t_{k+1}$, $0 \leq k \leq m_n - 1$, con $\lambda \in [0, 1]$ fijo, entonces

$$\sum_{k=0}^{m_n-1} (W(\tau_k^n) - W(t_k^n))^2 \xrightarrow[n \rightarrow \infty]{L^2(\Omega)} \lambda(b - a).$$

Lema 3.6. Sean \mathcal{P}^n particiones de $[0, T]$ como en (3.1) tales que $|\mathcal{P}^n| \rightarrow 0$, $n \rightarrow \infty$, y $\lambda \in [0, 1]$. Dados $\tau_k^n = (1 - \lambda)t_k + \lambda t_{k+1}$ consideramos la suma de Riemann

$$R_n := \sum_{k=0}^{m_n-1} W(\tau_k^n)(W(t_{k+1}^n) - W(t_k^n)). \quad (3.3)$$

Entonces

$$\lim_{n \rightarrow \infty} R_n \xrightarrow{L^2(\Omega)} \frac{W(T)^2}{2} + \left(\lambda - \frac{1}{2}\right)T. \quad (3.4)$$

Demostración. Para R_n dada por (3.3) con $W(0) = 0$, tenemos que

$$\begin{aligned} R_n &:= \frac{W(T)^2}{2} - \underbrace{\frac{1}{2} \sum_{k=0}^{m_n-1} (W(t_{k+1}^n) - W(t_k^n))^2}_{=:A} + \underbrace{\sum_{k=0}^{m_n-1} (W(\tau_k^n) - W(t_k^n))^2}_{=:B} \\ &\quad + \underbrace{\sum_{k=0}^{m_n-1} (W(t_{k+1}^n) - W(\tau_k^n))(W(\tau_k^n) - W(t_k^n))}_{=:C}. \end{aligned}$$

Por el Lema 3.2, $A \rightarrow \frac{T}{2}$ y $B \rightarrow \lambda T$ en $L^2(\Omega)$ si $n \rightarrow \infty$. Veamos qué ocurre con C . Nuevamente aplicando la normalidad e independencia de los incrementos de W

$$\begin{aligned} E[C^2] &= \sum_{k=0}^{m_n-1} E[(W(t_{k+1}^n) - W(\tau_k^n))^2] E[(W(\tau_k^n) - W(t_k^n))^2] = \sum_{k=0}^{m_n-1} (t_{k+1}^n - \tau_k^n)(\tau_k^n - t_k^n) \\ &= \sum_{k=0}^{m_n-1} (1 - \lambda)(t_{k+1}^n - t_k^n)\lambda(t_{k+1}^n - t_k^n) \leq \lambda(1 - \lambda)T|\mathcal{P}^n| \rightarrow 0. \end{aligned}$$

Por lo tanto, $C \rightarrow 0$ en $L^2(\Omega)$ si $n \rightarrow \infty$. \square

Observación 3.7. En particular, tomando $\lambda = 0$ (esto es $\tau_k^n = t_k^n$), se obtiene la integral de Itô:

$$\int_0^T W dW = \frac{W(T)^2}{2} - \frac{T}{2}. \quad (3.5)$$

3.2. Procesos estocásticos escalonados

Sea $W(\cdot)$ un proceso de Wiener unidimensional definido en el espacio de probabilidad (Ω, \mathcal{U}, P) . La σ -álgebra $\mathcal{W}(t) := \mathcal{U}(W(s) | 0 \leq s \leq t)$ es la historia del proceso de Wiener hasta el tiempo t (véase Definición 1.49).

Definición 3.8. La σ -álgebra $\mathcal{W}^+(t) := \mathcal{U}(W(s) - W(t) | s \geq t)$ es el futuro del proceso de Wiener mas allá del tiempo t .

Definición 3.9. Una familia $\mathcal{F}(\cdot)$ de σ -álgebras en \mathcal{U} se dice no anticipativa (o una filtración) con respecto al proceso de Wiener W si

- (a) $\mathcal{F}(t) \supseteq \mathcal{F}(s)$ para todo $t \geq s \geq 0$,
- (b) $\mathcal{F}(t) \supseteq \mathcal{W}(t)$ para todo $t \geq 0$,
- (c) $\mathcal{F}(t)$ es independiente de $\mathcal{W}^+(t)$ para todo $t \geq 0$.

Definición 3.10. Un proceso estocástico $G : [0, \infty) \times \Omega \rightarrow \mathbb{R}^n$ es no anticipativo con respecto a una filtración $\mathcal{F}(\cdot)$ si $G(t)$ es $\mathcal{F}(t)$ -medible, $\forall t \geq 0$, esto es, $\forall t \geq 0$ $(G(t))^{-1}(B) \in \mathcal{F}(t)$, para todo B conjunto medible Borel en \mathbb{R}^n .

Intuitivamente, el proceso estocástico G es no anticipativo para $t \geq 0$ si la variable aleatoria $G(t)$ depende solo de la información disponible en la σ -álgebra $\mathcal{F}(t)$.

Definición 3.11. Un proceso estocástico $G : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ se dice progresivamente medible si es no anticipativo y para todo $t \in [0, T]$ la aplicación

$$\begin{aligned} G : [0, t] \times \Omega &\longrightarrow \mathbb{R}^n \\ (s, \omega) &\longmapsto G(s, \omega) \end{aligned}$$

es $(\mathcal{B}_{[0,t]} \times \mathcal{F}(t))$ -medible, donde $\mathcal{B}_{[0,t]}$ es la σ -álgebra de Borel en $[0, t]$.

Intuitivamente, un proceso estocástico G es progresivamente medible si es no anticipativo y conjuntamente medible en las variables t y ω .

En lo que resta de esta sección 3.2 y en la sección 3.3 consideramos, por simplicidad en la exposición, procesos estocásticos valuados reales (esto es, $n = 1$). Al final de la sección 3.3 daremos unas breves indicaciones sobre la definición de la integral de Itô para procesos estocásticos n -dimensionales.

Definición 3.12. Denotamos por $\mathbb{L}^2(0, T)$ el conjunto de procesos estocásticos G en \mathbb{R} progresivamente medibles tales que $E \left[\int_0^T G^2 dt \right] < \infty$. Del mismo modo, $\mathbb{L}^1(0, T)$ en el conjunto de procesos estocásticos G en \mathbb{R} progresivamente medibles tales que $E \left[\int_0^T |G| dt \right] < \infty$.

Observación 3.13. Observar que si $X : [0, T] \times \Omega \rightarrow \mathbb{R}$, $E \left[\int_0^T X dt \right] = \int_{\Omega} \int_0^T X dt dP$.

Definición 3.14. Un proceso $G \in \mathbb{L}^2(0, T)$ se dice escalonado si existe una partición $\mathcal{P} = \{0 = t_0 < t_1 < \dots < t_m = T\}$ de $[0, T]$ tal que

$$G(t) = G_k, \quad \text{para } t_k \leq t < t_{k+1}, \tag{3.6}$$

donde G_k es una variable aleatoria $\mathcal{F}(t_k)$ -medible, para $k = 0, \dots, m - 1$.

Definición 3.15. Sea $G \in \mathbb{L}^2(0, T)$ un proceso escalonado (3.6). Se llama integral estocástica de Itô de G en el intervalo $[0, T]$ a

$$\int_0^T G dW := \sum_{k=0}^{m-1} G_k(W(t_{k+1}) - W(t_k)).$$

Se llama integral indefinida de Itô de G al proceso estocástico

$$\int_0^t G dW := \sum_{k=0}^{n-1} G_k(W(t_{k+1}) - W(t_k)) + G_n(W(t) - W(t_n)), \quad (3.7)$$

si $t \in (t_n, t_{n+1}]$, con $0 \leq n \leq m - 1$.

Observación 3.16. En particular, $\int_0^t G dW$ tiene trayectorias continuas para casi todo $w \in \Omega$, pues W verifica dicha propiedad (Teorema 2.12).

Lema 3.17. Sean $a, b \in \mathbb{R}$ y $G, H \in \mathbb{L}^2(0, T)$ procesos escalonados. Entonces:

$$i) \int_0^T (aG + bH) dW = a \int_0^T G dW + b \int_0^T H dW, \quad (3.8)$$

$$ii) E \left[\int_0^T G dW \right] = 0, \quad (3.9)$$

$$iii) E \left[\left(\int_0^T G dW \right)^2 \right] = E \left[\int_0^T G^2 dt \right]. \quad (3.10)$$

Demstración. (i) Es inmediato, pues podemos asumir sin pérdida de generalidad que G y H están definidos a partir de la misma partición \mathcal{P} . Bastaría aplicar la linealidad de una suma finita. Veamos (ii), para lo cual suponemos que $G(t) = G_k$ para $t_k \leq t < t_{k+1}$. Entonces

$$E \left[\int_0^T G dW \right] = \sum_{k=0}^{m-1} E[G_k(W(t_{k+1}) - W(t_k))].$$

Como G_k es $\mathcal{F}(t_k)$ -medible, $\mathcal{F}(t_k)$ es independiente de $W^+(t_k)$ y $W(t_{k+1}) - W(t_k)$ es $W^+(t_k)$ -medible, entonces G_k es independiente de $W(t_{k+1}) - W(t_k)$. Por tanto,

$$E[G_k(W(t_{k+1}) - W(t_k))] = E[G_k]E[W(t_{k+1}) - W(t_k)] = 0.$$

Respecto a (iii), tenemos

$$E \left[\left(\int_0^T G dW \right)^2 \right] = \sum_{j,k=0}^{m-1} E[G_k G_j (W(t_{k+1}) - W(t_k))(W(t_{j+1}) - W(t_j))].$$

Por el mismo razonamiento anterior, si $j < k$, $W(t_{k+1}) - W(t_k)$ es independiente de $G_k G_j (W(t_{j+1}) - W(t_j))$, con $E[W(t_{k+1}) - W(t_k)] = 0$. En consecuencia,

$$\begin{aligned} E \left[\left(\int_0^T G dW \right)^2 \right] &= \sum_{k=0}^{m-1} E[G_k^2] E[(W(t_{k+1}) - W(t_k))^2] = \sum_{k=0}^{m-1} E[G_k^2] (t_{k+1} - t_k) \\ &= \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} E[G_k^2] dt = E \left[\sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} G_k^2 dt \right] = E \left[\int_0^T G^2 dt \right]. \quad \square \end{aligned}$$

Finalizamos esta sección con una serie de resultados de aproximación de un proceso $G \in \mathbb{L}^2(0, T)$ mediante procesos escalonados. Este argumento permitirá definir la integral de Itô de G mediante paso al límite.

Teorema 3.18. *Sea $G \in \mathbb{L}^2(0, T)$. Existe una sucesión de procesos escalonados acotados $G^n \in \mathbb{L}^2(0, T)$ tal que*

$$\lim_{n \rightarrow \infty} E \left[\int_0^T |G - G^n|^2 dt \right] \rightarrow 0. \quad (3.11)$$

Este teorema es crucial para definir la integral de Itô $\int_0^T G dW$ en la sección 3.3. Para su demostración, que posponemos al final de esta sección 3.2, haremos uso de los tres siguientes lemas ([10, Cap. 3], [11, Lec. 15]).

Lema 3.19. *Sea $G \in \mathbb{L}^2(0, T)$ un proceso continuo casi seguro verificando que existe M tal que $P \left(\left\{ w \in \Omega : \sup_{0 \leq t \leq T} |G_t(w)| \leq M \right\} \right) = 1$. Entonces existe una sucesión de procesos escalonados $G^n \in \mathbb{L}^2(0, T)$ cumpliendo (3.11).*

Demostración. En primer lugar, consideramos una sucesión de particiones $\mathcal{P} = \{t_j^n\}_{j=0}^{m_n}$ de $[0, T]$ tal que $|\mathcal{P}| \rightarrow 0, n \rightarrow \infty$. Ahora, dado el proceso G definimos el proceso G^n como $G_t^n = G_{t_j^n}$ para todo $t \in [t_j^n, t_{j+1}^n)$, de modo que G^n es un proceso escalonado y además progresivamente medible. Ya que G es continuo casi seguro, tenemos que $G(w) = \lim_{n \rightarrow \infty} G^n(w)$, casi seguro. Luego, la sucesión de procesos progresivamente medibles $\{G^n\}_{n \geq 0}$ converge casi seguro a G . Por hipótesis, también tenemos que $P \left(\left\{ w \in \Omega : \sup_{0 \leq t \leq T} |G_t(w)| \leq M \right\} \right) = 1$. Así, el Teorema de la Convergencia Dominada de Lebesgue [12, p. 26], junto con la convergencia casi segura permite escribir

$$\lim_{n \rightarrow \infty} E \left[\int_0^T (G^n - G)^2 dt \right] = \lim_{n \rightarrow \infty} \int_{\Omega} \int_0^T (G^n - G)^2 dt dP = 0. \quad \square$$

Lema 3.20. *Sea $G \in \mathbb{L}^2(0, T)$ un proceso acotado casi seguro, esto es $|G| \leq M$, c.s. Existe una sucesión de procesos continuos acotados $G^n \in \mathbb{L}^2(0, T)$ cumpliendo (3.11).*

Demostración. Sea el proceso G^n definido como $G_t^n = n \int_{t-\frac{1}{n}}^t G_s ds$. Tenemos que $|G^n| \leq n(\frac{1}{n})M = M$. Por otro lado, $G_{t'}^n - G_t^n = n \left(\int_t^{t'} X_s ds - \int_{t-\frac{1}{n}}^{t'-\frac{1}{n}} X_s ds \right)$. Luego, $|G_{t'}^n - G_t^n| \leq 2nM|t' - t|$.

Como G es integrable Riemann casi seguro, G^n es continuo y acotado casi seguro. Por ello, para casi todo ω , el conjunto de puntos de discontinuidad de $G(w)$ tiene medida cero y, por el Teorema Fundamental del Cálculo [12, p. 140], para todo punto de continuidad t tenemos que $\lim_{n \rightarrow \infty} G^n(w) = G(w)$. Luego concluimos que G^n converge puntualmente a G casi seguro. Nuevamente, aplicando el Teorema de la Convergencia Dominada [12, p. 26] se obtiene el resultado. \square

Lema 3.21. *Sea $G \in \mathbb{L}^2(0, T)$. Existe una sucesión de procesos acotados casi seguro $G^n \in \mathbb{L}^2(0, T)$ cumpliendo (3.11).*

Demostración. Definimos G^n como

$$G^n = \begin{cases} G, & -n \leq G \leq n, \\ -n, & G < -n, \\ n, & G > n. \end{cases}$$

Tenemos que $G^n \rightarrow G, n \rightarrow \infty$, casi seguro con respecto a $\omega \in \Omega$ y $t \in [0, T]$, y $|G^n| = \min\{n, |G|\}$, para cada n . Puesto que $(a - b)^2 \leq 2(a^2 + b^2)$ se deduce que

$$\int_0^T (G^n - G)^2 dt \leq 2 \int_0^T (G^n)^2 dt + 2 \int_0^T G^2 dt \leq 4 \int_0^T G^2 dt.$$

Como $E[\int_0^T G^2 dt] < \infty$, se obtiene el resultado a partir del Teorema de la Convergencia Dominada [12, p.26]. \square

Demostración. (Teorema 3.18). Es consecuencia de combinar los lemas 3.19, 3.20 y 3.21. \square

Observación 3.22. Podemos enunciar un resultado similar al Teorema 3.18 para procesos $F \in \mathbb{L}^1(0, T)$.

Teorema 3.23. *Sea $F \in \mathbb{L}^1(0, T)$. Existe $\{F^n\}_{n \geq 1} \subset \mathbb{L}^1(0, T)$ sucesión de procesos escalonados acotados tales que*

$$\lim_{n \rightarrow \infty} E \left[\int_0^T |F^n - F| dt \right] \rightarrow 0. \quad (3.12)$$

Demostración. Basta aplicar el Teorema 3.18 a $\sqrt{F^+}$ y $\sqrt{F^-}$, con la descomposición usual $F = \sqrt{F^+} - F^-$, teniendo en cuenta que $F^+, \sqrt{F^-} \in L^2(0, T)$, así como la desigualdad de Cauchy-Schwartz en $L^2(0, T)$ y $L^2(\Omega)$. \square

3.3. La integral de Itô

Definición 3.24. *Sea $G \in L^2(0, T)$ y $\{G^n\}_{n=1}^\infty \subset L^2(0, T)$ una sucesión de procesos escalonados acotados para cada $n \geq 1$ verificando (3.11). Se llama integral de Itô de G*

$$\int_0^T G dW := \lim_{n \rightarrow \infty} \int_0^T G^n dW, \quad \text{en } L^2(\Omega). \quad (3.13)$$

Lema 3.25. *Si $G \in L^2(0, T)$ entonces $\int_0^T G dW$ está bien definida.*

Demostración. 1. Veamos que en efecto $\int_0^T G dW \in L^2(\Omega)$.

Sea $\mathcal{G}^n := \int_0^T G^n dW$. $\mathcal{G}^n \in L^2(\Omega)$ pues $E[(\mathcal{G}^n)^2] = E[\int_0^T (G^n)^2 dt] < \infty$ en virtud de (3.10). Ahora, usando que $(a - b)^2 \leq 2(a^2 + b^2)$, se tiene que por (3.10) que

$$\begin{aligned} \|\mathcal{G}^n - \mathcal{G}^m\|_{L^2(\Omega)}^2 &= E \left[\left(\int_0^T (G^n - G^m) dW \right)^2 \right] = E \left[\int_0^T (G^n - G^m)^2 dt \right] \\ &\leq 2E \left[\int_0^T (G^n - G)^2 dt \right] + 2E \left[\int_0^T (G^m - G)^2 dt \right] \rightarrow 0, \quad \text{si } n, m \rightarrow \infty. \end{aligned}$$

Luego, $\{\mathcal{G}^n\}_{n=1}^\infty$ es una sucesión de Cauchy en $(L^2(\Omega), \|\cdot\|_{L^2(\Omega)})$ y como este espacio es completo (Observación 1.22), existe $\lim_{n \rightarrow \infty} \mathcal{G}^n = \lim_{n \rightarrow \infty} \int_0^T G^n dW \in L^2(\Omega)$.

2. El mismo argumento muestra que si $\{\hat{G}^n\}_{n=1}^\infty \subset L^2(0, T)$ es otra sucesión de procesos escalonados acotados verificando (3.11), entonces para $\hat{\mathcal{G}}^n := \int_0^T \hat{G}^n dW$, se cumple que $\|\mathcal{G}^n - \hat{\mathcal{G}}^n\|_{L^2(\Omega)} \rightarrow 0$, si $n \rightarrow \infty$. Por tanto, $\lim_{n \rightarrow \infty} \int_0^T \hat{G}^n dW = \lim_{n \rightarrow \infty} \int_0^T G^n dW$. \square

Observación 3.26. En particular, teniendo en cuenta la demostración del Lema (3.19), si $G \in \mathbb{L}^2(0, T)$ y $t \mapsto G(t, \omega)$ es continua para casi todo ω , entonces

$$\sum_{k=0}^{m_n-1} G(t_k^n)(W(t_{k+1}^n) - W(t_k^n)) \xrightarrow[n \rightarrow \infty]{L^2(\Omega)} \int_0^T G dW,$$

donde $\mathcal{P}^n := \{0 = t^0 < \dots < t_{m_n}^n = T\}$ es cualquier sucesión de particiones con $|\mathcal{P}^n| \rightarrow 0$. Esto muestra la consistencia de la definición de la integral de Itô con la propiedad enunciada en (3.5).

Teorema 3.27. (*Propiedades de la integral de Itô*). Sean $a, b \in \mathbb{R}$ y $G, H \in \mathbb{L}^2(0, T)$.

$$(i) \int_0^T (aG + bH)dW = a \int_0^T GdW + b \int_0^T HdW, \quad (3.14)$$

$$(ii) E \left[\int_0^T GdW \right] = 0, \quad (3.15)$$

$$(iii) E \left[\left(\int_0^T GdW \right)^2 \right] = E \left[\int_0^T G^2 dt \right], \quad (\text{isometría de Itô}) \quad (3.16)$$

$$(iv) E \left[\left(\int_0^T GdW \right) \left(\int_0^T HdW \right) \right] = E \left[\int_0^T GH dt \right]. \quad (3.17)$$

Demostración. (i) Es obvio por linealidad del límite y por la correspondiente propiedad para procesos escalonados.

(ii) Sea $\mathcal{G} := \int_0^T GdW$, $\{\mathcal{G}^n\}_{n \geq 1} \subset \mathbb{L}^2(0, T)$ es una sucesión de procesos escalonados acotados verificando (3.11) y $\mathcal{G}^n = \int_0^T G^n dW$. Entonces, usando (3.9)

$$|E[\mathcal{G}]| \leq |E[\mathcal{G} - \mathcal{G}^n]| + |E[\mathcal{G}^n]| = |E[\mathcal{G} - \mathcal{G}^n]|.$$

Luego, por la desigualdad de Cauchy-Schwartz (1.2).

$$|E[\mathcal{G}]| \leq E[(\mathcal{G} - \mathcal{G}^n)^2]^{\frac{1}{2}} \rightarrow 0, \text{ si } n \rightarrow \infty. \text{ Luego } E[\mathcal{G}] = 0.$$

(iii) Tenemos que $E[\mathcal{G}^2] = E[(\mathcal{G} - \mathcal{G}^n)^2] + 2E[(\mathcal{G} - \mathcal{G}^n)\mathcal{G}^n] + E[(\mathcal{G}^n)^2]$. Ahora $E[(\mathcal{G} - \mathcal{G}^n)^2] \rightarrow 0$, $n \rightarrow \infty$, y por desigualdad de Cauchy-Schwartz

$$|E[(\mathcal{G} - \mathcal{G}^n)\mathcal{G}^n]| \leq E[(\mathcal{G} - \mathcal{G}^n)^2]^{\frac{1}{2}} E[(\mathcal{G}^n)^2]^{\frac{1}{2}} \rightarrow 0, \text{ } n \rightarrow \infty.$$

Luego, por (3.10), $E[\mathcal{G}^2] = \lim_{n \rightarrow \infty} E[(\mathcal{G}^n)^2] = \lim_{n \rightarrow \infty} E \left[\int_0^T (G^n)^2 dt \right]$. Por otra parte,

$$\begin{aligned} \left| E \left[\int_0^T (G^n)^2 dt \right] - E \left[\int_0^T G^2 dt \right] \right| &= \left| E \left[\int_0^T (G^n - G)(G^n + G) dt \right] \right| \\ &\leq E \left[\int_0^T |G^n - G| |G^n + G| dt \right] \\ &\leq \left(E \left[\int_0^T (G^n - G)^2 dt \right] \right)^{\frac{1}{2}} \left(E \left[\int_0^T (G^n + G)^2 dt \right] \right)^{\frac{1}{2}}, \end{aligned}$$

donde en la última línea hemos aplicado la desigualdad de Cauchy-Schwartz en $L^2([0, T])$ y $L^2(\Omega)$, respectivamente. En virtud del Teorema 3.18 se tiene que $\lim_{n \rightarrow \infty} E \left[\int_0^T (G^n)^2 dt \right] = E \left[\int_0^T G^2 dt \right]$.

(iv) Teniendo en cuenta que $2ab = (a + b)^2 - a^2 - b^2$, si $\mathcal{G} = \int_0^T G dW$ y $\mathcal{H} = \int_0^T H dW$, entonces aplicando (iii)

$$\begin{aligned} 2E[\mathcal{G}\mathcal{H}] &= E[(\mathcal{G} + \mathcal{H})^2] - E[\mathcal{G}^2] - E[\mathcal{H}^2] \\ &= E\left[\int_0^T (G + H)^2 dt\right] - E\left[\int_0^T G^2 dt\right] - E\left[\int_0^T H^2 dt\right] \\ &= 2E\left[\int_0^T GH dt\right]. \quad \square \end{aligned}$$

Finalizamos esta sección introduciendo el concepto de la integral indefinida de Itô y la extensión de la integral de Itô a procesos estocásticos n -dimensionales.

Definición 3.28. Sea $G \in \mathbb{L}^2(0, T)$. Se llama *integral indefinida de Itô de G* al proceso estocástico

$$I(t) := \int_0^t G dW, \quad 0 \leq t \leq T.$$

Teorema 3.29. ([8, Teoremas 3.2.5 y 3.2.6]). Sea $G \in \mathbb{L}^2(0, T)$. Entonces la integral indefinida de G , $I(\cdot)$, es una martingala. Además, $I(\cdot)$ tiene una versión con trayectorias continuas casi seguro.

Definición 3.30. Un proceso estocástico $\mathbf{G} = (G^1, G^2, \dots, G^n)$ en \mathbb{R}^n pertenece a $\mathbb{L}_n^p(0, T)$ si $G^i \in \mathbb{L}^p(0, T)$, $i = 1, \dots, n$, con $p = 1, 2$.

Definición 3.31. Sea $\mathbb{G} \in \mathbb{L}_n^2(0, T)$. Se llama *integral de Itô de \mathbb{G}* a la variable aleatoria en \mathbb{R}^n dada por $\int_0^T \mathbf{G} dW = \left(\int_0^T G^i dW\right)_{i=1}^n$.

3.4. Diferenciales estocásticas: reglas de la cadena y del producto de Itô

Definición 3.32. Sean $F \in \mathbb{L}^1(0, T)$, $G \in \mathbb{L}^2(0, T)$ y W un proceso de Wiener unidimensional. Un proceso estocástico X en \mathbb{R} tiene la diferencial estocástica

$$dX = F dt + G dW \tag{3.18}$$

si $X(r) = X(s) + \int_s^r F dt + \int_s^r G dW$ casi seguro, para $0 \leq s \leq r \leq T$.

Lema 3.33. Sea W un proceso de Wiener unidimensional. Entonces

$$(i) d(W^2) = 2W dW + dt, \quad (ii) d(tW) = W dt + t dW.$$

Demostración. (i) Es consecuencia inmediata del Lema 3.6 con $\lambda = 0$. Respecto a (ii), dados $r, s, 0 \leq s \leq r \leq T$, nótese que

$$\int_s^r t dW = \lim_{n \rightarrow \infty} \sum_{k=0}^{m_n-1} t_k^n (W(t_{k+1}^n) - W(t_k^n)) \quad \text{en } L^2(\Omega), \tag{3.19}$$

donde $\mathcal{P}^n = \{s = t_0^n < t_1^n < \dots < t_{m_n}^n = r\}$ es una sucesión de particiones de $[s, r]$, con $|\mathcal{P}^n| \rightarrow 0$. De manera similar, como $t \mapsto W(t)$ es continua casi seguro, tenemos para la integral de Riemann

$$\int_s^r W dt = \lim_{n \rightarrow \infty} \sum_{k=0}^{m_n-1} W(t_{k+1}^n)(t_{k+1}^n - t_k^n), \quad \text{casi seguro.} \quad (3.20)$$

Por tanto, sumando (3.19) y (3.20), y considerando que en (3.19) la convergencia también es casi seguro tal vez mediante subsucesiones, entonces $\int_s^r t dW + \int_s^r W dt = rW(r) - sW(s)$, casi seguro. \square

El principal resultado de esta sección es la regla de la cadena de Itô que queda establecida en el siguiente resultado. Posponemos su demostración al final de esta sección.

Teorema 3.34. (Regla de la cadena de Itô). *Sea X un proceso estocástico en \mathbb{R} con diferencial estocástica $dX = Fdt + GdW$, donde $F \in \mathbb{L}^1(0, T)$ y $G \in \mathbb{L}^2(0, T)$. Sea $u : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$, continua tal que sus derivadas $u_t = \frac{\partial u}{\partial t}$, $u_x = \frac{\partial u}{\partial x}$ y $u_{xx} = \frac{\partial^2 u}{\partial x^2}$ existen y son continuas. Entonces $Y := u(t, X)$ tiene la diferencial estocástica*

$$dY = u_t dt + u_x dX + \frac{1}{2} u_{xx} G^2 dt = (u_t + u_x F + \frac{1}{2} u_{xx} G^2) dt + u_x G dW, \quad (3.21)$$

donde las derivadas de u están evaluadas en $(t, X(t))$.

Observación 3.35. (i) La fórmula de Itô indica que para $0 \leq s \leq r \leq T$ se tiene que

$$\begin{aligned} Y(r) - Y(s) &= u(r, X(r)) - u(s, X(s)) \\ &= \int_s^r [u_t(t, X(t)) + u_x(t, X(t))F + \frac{1}{2} u_{xx}(t, X(t))G^2] dt \\ &\quad + \int_s^r u_x(t, X(t))G dW, \quad \text{casi seguro.} \end{aligned}$$

(ii) Como $X(t) = X(0) + \int_0^t F ds + \int_0^t G dW$, X tiene trayectorias continuas casi seguro (ver Teorema 3.29). Así para casi todo ω , las funciones $t \mapsto u_t(t, X(t))$, $u_x(t, X(t))$, $u_{xx}(t, X(t))$ son continuas y, en conclusión, las integrales anteriores están bien definidas.

Ejemplo 3.36. Sea $X = W$ e $Y = X^m$. Entonces $dX = dW$, con $F \equiv 0$, $G \equiv 1$. Utilizando la fórmula de Itô con $u(t, X) = X^m$, sigue que

$$d(Y) = d(W^m) = mW^{m-1} dW + \frac{1}{2} m(m-1)W^{m-2} dt.$$

En particular, si $m = 2$, $d(W^2) = 2W dW + dt$, tal como ya sabemos.

Ejemplo 3.37. Sea $X = W$, $Y = u(t, X)$, con $u(t, X) = e^{\lambda x - \frac{\lambda^2 t}{2}}$. Entonces por la fórmula de Itô, con $F \equiv 0$ y $G \equiv 1$:

$$dY = \left(-\frac{\lambda^2}{2} Y + \lambda Y \cdot 0 + \frac{\lambda^2}{2} Y \cdot 1^2\right) dt + \lambda Y \cdot 1 dW = \lambda Y dW.$$

Así, puesto que $X(0) = W(0) = 1$, el proceso estocástico Y verifica $\begin{cases} dY = \lambda Y dW \\ Y(0) = 1 \end{cases}$.

Teorema 3.38. (Fórmula del producto de Itô). *Sean X_i , $i = 1, 2$, procesos estocásticos con diferenciales $dX_i = F_i dt + G_i dW$, con $F_i \in \mathbb{L}^1(0, T)$, $G_i \in \mathbb{L}^2(0, T)$, $i = 1, 2$. Entonces*

$$d(X_1 X_2) = X_1 dX_2 + X_2 dX_1 + G_1 G_2 dt. \quad (3.22)$$

Observación 3.39. La versión integral de (3.22) representa la regla de integración por partes de Itô. El término $G_1 G_2 dt$ representa el término de corrección de Itô.

Demostración. En primer lugar, considerando la regla de la cadena de Itô para $d(X_i^2)$, $i = 1, 2$, sigue con $u(t, x) = x^2$, que

$$d(X_i^2) = 2X_i dX_i + G_i^2 dt, \quad i = 1, 2. \quad (3.23)$$

Ahora, como $ab = \frac{1}{2}[(a+b)^2 - a^2 - b^2]$, entonces por linealidad de la integral:

$$d(X_1 X_2) = \frac{1}{2}d((X_1 + X_2)^2) - \frac{1}{2}d(X_1^2) - \frac{1}{2}d(X_2^2), \quad (3.24)$$

y aplicando (3.23)

$$\begin{aligned} d(X_1 X_2) &= \frac{1}{2}[2(X_1 + X_2)d(X_1 + X_2) + (G_1 + G_2)^2 dt] \\ &\quad - \frac{1}{2}[2X_1 dX_1 + G_1^2 dt] - \frac{1}{2}[2X_2 dX_2 + G_2^2 dt] \\ &= X_1 dX_2 + X_2 dX_1 + G_1 G_2 dt. \quad \square \end{aligned}$$

Observación 3.40. Una forma heurística de deducir la regla del producto y de la cadena de Itô se basa en considerar desarrollo en serie de Taylor con las identificaciones

$$(dW)^2 = dt, \quad dt dW = 0, \quad (dt)^2 = 0, \quad (3.25)$$

que guardan relación con el Lema 3.2 1) Si $f = f(x_1, x_2)$:

$$\Delta f = f_{x_1} \Delta x_1 + f_{x_2} \Delta x_2 + \frac{1}{2}\{f_{x_1 x_1} (\Delta x_1)^2 + 2f_{x_1 x_2} (\Delta x_1)(\Delta x_2) + f_{x_2 x_2} (\Delta x_2)^2\} + \mathcal{O}(\|(\Delta x_1, \Delta x_2)\|^3).$$

Tomando $f(x_1, x_2) = x_1 x_2$

$$\Delta f = x_2 \Delta x_1 + x_1 \Delta x_2 + (\Delta x_1)(\Delta x_2) + \mathcal{O}(\|(\Delta x_1, \Delta x_2)\|^3)$$

Si $dX_i = F_i dt + G_i dW$, $i = 1, 2$, tendremos de (3.25)

$$\begin{aligned} d(X_1 X_2) &= X_2 dX_1 + X_1 dX_2 + dX_1 dX_2 \\ &= X_2 dX_1 + X_1 dX_2 + G_1 G_2 dt. \end{aligned}$$

2) Si $dX = F dt + G dW$ e $Y = u(t, X)$, u regular, entonces

$$\Delta Y = \Delta u = u_t \Delta t + u_x \Delta x + \frac{1}{2}\{u_{xx} (\Delta x)^2 + 2u_{tx} (\Delta t)(\Delta x) + u_{tt} (\Delta t)^2\} + \mathcal{O}(\|(\Delta t, \Delta x)\|^3).$$

Nuevamente, usando las identificaciones (3.25) :

$$\begin{aligned} dY &= u_t dt + u_x dX + \frac{1}{2}u_{xx} (dX)^2 \\ &= u_t dt + u_x dX + \frac{1}{2}u_{xx} G^2 dt. \end{aligned}$$

Demostración. (Teorema 3.34). Tenemos que $X(t) = X(0) + \int_0^t F ds + \int_0^t G dW$ casi seguro y en particular X tiene trayectorias continuas casi seguro (Teorema 3.29). Tenemos que probar que

$$Y(t) = Y(0) + \int_0^t [u_t(s, X(s)) + u_x(s, X(s))F + \frac{1}{2}u_{xx}(s, X(s))G^2] ds + \int_0^t u_x(s, X(s))G dW, \quad \text{c.s.} \quad (3.26)$$

1°) Suponemos inicialmente que u y sus derivadas parciales están acotadas en $[0, T] \times \mathbb{R}$, y que F y G son procesos escalonados acotados. Podemos asumir entonces además

que F y G son procesos constantes para $t \in [0, T]$ (pues bastaría probar (3.26) en subintervalos). Luego, $X(t) = X(0) + Ft + GW(t)$, $t \in [0, T]$, con $E[|F|] < \infty$, $E[G^2] < \infty$.

Para probar (3.26), fijamos $t \in [0, T]$ y consideramos particiones arbitrarias $\mathcal{P} = \{t_j\}_{j=0}^n$ de $[0, T]$, con $|\mathcal{P}^n| \rightarrow 0$, si $n \rightarrow \infty$ (donde por simplicidad de escritura suprimimos la dependencia de los nodos t_j de n). Denotamos

$$\Delta t_j = t_{j+1} - t_j, \quad X_j = X(t_j), \quad \Delta X_j = X(t_{j+1}) - X(t_j).$$

Entonces, usando desarrollo de Taylor y el Teorema del valor medio, tenemos casi seguro que

$$\begin{aligned} Y(t) - Y(0) &= u(t, X(t)) - u(0, X(0)) = \sum_{j=0}^{n-1} (u(t_{j+1}, X_{j+1}) - u(t_j, X_j)) \\ &= \sum_{j=0}^{n-1} ((u(t_{j+1}, X_{j+1}) - u(t_j, X_{j+1})) + (u(t_j, X_{j+1}) - u(t_j, X_j))) \\ &= \sum_{j=0}^{n-1} \{u_t(t_j + \theta_j^{(1)} \Delta t_j, X_{j+1}) \Delta t_j + u_x(t_j, X_j) \Delta X_j + \frac{1}{2} u_{xx}(t_j, X_j + \theta_j^{(2)} \Delta X_j) (\Delta X_j)^2\}, \end{aligned}$$

con $0 \leq \theta_j^{(1)}, \theta_j^{(2)} \leq 1$, para todo $j = 0, \dots, n-1$. Denotando ahora $\hat{t}_j = t_j + \theta_j^{(1)} \Delta t_j$, $\hat{X}_j = X_j + \theta_j^{(2)} \Delta X_j$, y teniendo en cuenta que $\Delta X_j = F \Delta t_j + G \Delta W_j$, con $\Delta W_j = W(t_{j+1}) - W(t_j)$, sigue que

$$\begin{aligned} Y(t) - Y(0) &= \sum_{j=0}^{n-1} \underbrace{u_t(\hat{t}_j, X_{j+1}) \Delta t_j + u_x(t_j, X_j) F \Delta t_j + \frac{1}{2} u_{xx}(t_j, X_j) G^2 \Delta t_j}_{:=A_j} \\ &\quad + \underbrace{u_x(t_j, X_j) G \Delta W_j}_{:=B_j} + \frac{1}{2} \underbrace{[u_{xx}(t_j, \hat{X}_j) - u_{xx}(t_j, X_j)] (\Delta X_j)^2}_{:=C_j} \\ &\quad + \frac{1}{2} \underbrace{u_{xx}(t_j, X_j) [(\Delta X_j)^2 - G^2 (\Delta t_j)]}_{:=D_j}. \end{aligned}$$

Tenemos así que $Y(t) - Y(0) = S_{n,1} + S_{n,2} + S_{n,3} + S_{n,4}$, con

$$S_{n,1} := \sum_{j=0}^{n-1} A_j, \quad S_{n,2} := \sum_{j=0}^{n-1} B_j, \quad S_{n,3} := \sum_{j=0}^{n-1} C_j, \quad S_{n,4} := \sum_{j=0}^{n-1} D_j.$$

i) Analicemos $S_{n,1}$. Puesto que u_t , u_x , u_{xx} son continuas y acotadas, tenemos c.s. que

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_{0 \leq j \leq n-1} \sup_{s \in [t_j, t_{j+1}]} |u_t(\hat{t}_j, X_{j+1}) - u_t(s, X(s))| &= 0, \\ \lim_{n \rightarrow \infty} \sup_{0 \leq j \leq n-1} \sup_{s \in [t_j, t_{j+1}]} |u_x(t_j, X_j) - u_x(s, X(s))| &= 0, \\ \lim_{n \rightarrow \infty} \sup_{0 \leq j \leq n-1} \sup_{s \in [t_j, t_{j+1}]} |u_{xx}(t_j, X_j) - u_{xx}(s, X(s))| &= 0. \end{aligned}$$

Luego $\lim_{n \rightarrow \infty} S_{n,1} = \int_0^t [u_t(s, X(s)) + u_x(s, X(s))F + \frac{1}{2} u_{xx}(s, X(s))G^2] ds$, c.s.

ii) Analicemos ahora $S_{n,2}$. Por definición de integral de Itô:

$$\lim_{n \rightarrow \infty} S_{n,2} = \int_0^t u_x(s, X(s)) G dW \quad \text{en } L^2(\Omega).$$

iii) Respecto a $S_{n,3}$, puesto que $(\Delta X_j)^2 = F^2(\Delta t_j)^2 + 2FG(\Delta t_j)(\Delta W_j) + G^2(\Delta W_j)^2$, sigue que

$$\sum_{j=0}^{n-1} (\Delta X_j)^2 \leq F^2(\max_{0 \leq j \leq n-1} \Delta t_j)t + 2FG(\max_{0 \leq j \leq n-1} \Delta t_j)W(t) + G^2 \sum_{j=0}^{n-1} (\Delta W_j)^2,$$

donde los dos primeros sumandos tienden a cero si $n \rightarrow \infty$ casi seguro, mientras que el tercer sumando tiende a $G^2 t$ en $L^2(\Omega)$ en virtud de (3.2). Puesto que convergencia casi seguro y en media cuadrática implican convergencia en probabilidad (Teorema 1.29),

sigue que $\lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} (\Delta X_j)^2 = G^2 t$, en probabilidad.

Como $\lim_{n \rightarrow \infty} \sup_{0 \leq j \leq n-1} |u_{xx}(t_j, \hat{X}_{j+1}) - u_{xx}(t_j, X_j)| = 0$ casi seguro, se tiene que $\lim_{n \rightarrow \infty} S_{n,3} = 0$, en probabilidad.

iv) Para estudiar $S_{n,4}$, como $(\Delta X_j)^2 - G^2(\Delta t_j) = F^2(\Delta t_j)^2 + 2FG(\Delta t_j)(\Delta W_j) + G^2[(\Delta W_j)^2 - (\Delta t_j)]$, sigue que $S_{n,4} = S_{n,4}^{(1)} + S_{n,4}^{(2)} + S_{n,4}^{(3)}$, con

$$S_{n,4}^{(1)} = \frac{1}{2} F^2 \sum_{j=0}^{n-1} u_{xx}(t_j, X_j) (\Delta t_j)^2,$$

$$S_{n,4}^{(2)} = FG \sum_{j=0}^{n-1} u_{xx}(t_j, X_j) (\Delta t_j) (\Delta W_j),$$

$$S_{n,4}^{(3)} = \frac{1}{2} G^2 \sum_{j=0}^{n-1} u_{xx}(t_j, X_j) [(\Delta W_j)^2 - (\Delta t_j)].$$

Como en (iii), tendremos que $\lim_{n \rightarrow \infty} S_{n,4}^{(1)} = \lim_{n \rightarrow \infty} S_{n,4}^{(2)} = 0$ casi seguro. Veamos ahora que $S_{n,4}^{(3)} \rightarrow 0$ en $L^2(\Omega)$. En primer lugar, observemos que las variables aleatorias $(\Delta W_j)^2 - (\Delta t_j)$ son independientes, $0 \leq j \leq n-1$, y

$$E[(\Delta W_j)^2 - (\Delta t_j)] = 0, \quad E[((\Delta W_j)^2 - (\Delta t_j))^2] = 2(\Delta t_j)^2, \quad 0 \leq j \leq n-1.$$

Luego usando independencia y acotación de u_{xx}

$$\begin{aligned} E \left[\left(\sum_{j=0}^{n-1} u_{xx}(t_j, X_j) [(\Delta W_j)^2 - (\Delta t_j)] \right)^2 \right] &= \\ &= \sum_{i,j=0}^{n-1} E[u_{xx}(t_i, X_i) u_{xx}(t_j, X_j) ((\Delta W_i)^2 - (\Delta t_i)) ((\Delta W_j)^2 - (\Delta t_j))] \\ &= \sum_{j=0}^{n-1} E[u_{xx}(t_j, X_j)^2] E[((\Delta W_j)^2 - (\Delta t_j))^2] \\ &\leq 2C \sum_{j=0}^{n-1} (\Delta t_j)^2 \leq 2Ct \max_{0 \leq j \leq n-1} (\Delta t_j) \rightarrow 0, \quad \text{si } n \rightarrow \infty. \end{aligned}$$

Por tanto, $E[(S_{n,4}^{(3)})^2] \rightarrow 0$, si $n \rightarrow \infty$.

Finalmente, puesto que convergencia casi seguro y en media cuadrática implican convergencia en probabilidad, y esta convergencia implica convergencia casi seguro con respecto a alguna subsucesión $\{n_k\}_{k=1}^\infty$ (Teorema 1.29) tenemos casi seguro que

$$\begin{aligned} Y(t) - Y(0) &= \lim_{k \rightarrow \infty} \sum_{j=0}^{n_k-1} u_t(\hat{t}_j, X_{j+1}) \Delta t_j + u_x(t_j, X_j) F \Delta t_j + \frac{1}{2} u_{xx}(t_j, X_j) G^2 \Delta t_j \\ &\quad + u_x(t_j, X_j) G \Delta W_j + \frac{1}{2} [u_{xx}(t_j, \hat{X}_j) - u_{xx}(t_j, X_j)] (\Delta X_j)^2 \\ &\quad + \frac{1}{2} u_{xx}(t_j, X_j) [(\Delta X_j)^2 - G^2(\Delta t_j)] \\ &= \int_0^t [u_t(s, X(s)) + u_x(s, X(s)) F + \frac{1}{2} G^2 u_{xx}(s, X(s))] ds + \int_0^t u_x(s, X(s)) G dW. \end{aligned}$$

2º) En general, si $F \in \mathbb{L}^1(0, T)$ y $G \in \mathbb{L}^2(0, T)$, por los Teoremas 3.18 y 3.23, existen sucesiones $\{F^n\}_{n \geq 1} \subset \mathbb{L}^1(0, T)$ y $\{G^n\}_{n \geq 1} \subset \mathbb{L}^2(0, T)$ de procesos escalonados acotados verificando (3.12) y (3.11), respectivamente. Además, respecto a u, u_t, u_x, u_{xx} , éstas se pueden aproximar por funciones $u_n, (u_n)_t, (u_n)_x, (u_n)_{xx}$ acotadas con convergencia uniforme en compactos de $[0, T] \times \mathbb{R}$. La convergencia uniforme junto con (3.11)-(3.12) permiten demostrar (3.26) en general mediante paso al límite (ver [2, p. 196-200] para más detalles). \square

Finalizamos este capítulo mencionando que la definición de diferencial estocástica, así como la regla de la cadena de Itô, se pueden extender al caso n -dimensional del modo siguiente.

Definición 3.41. Sean $F \in \mathbb{L}_n^1(0, T)$, $G \in \mathbb{L}_n^2(0, T)$ y W un proceso de Wiener unidimensional. Un proceso estocástico $X = (X^1, \dots, X^n)$ en \mathbb{R}^n tiene la diferencial estocástica $dX = Fdt + GdW$ si $X(r) = X(s) + \int_s^r Fdt + \int_s^r GdW$ casi seguro, para $0 \leq s \leq r \leq T$. En particular, $dX^i = F^i dt + G^i dW$, $i = 1, \dots, n$.

Teorema 3.42. (Regla de la cadena de Itô generalizada). Sean X^i procesos estocásticos en \mathbb{R} con diferenciales estocásticas $dX^i = F^i dt + G^i dW$, ($i = 1, \dots, n$), con $F^i \in \mathbb{L}^1(0, T)$ y $G^i \in \mathbb{L}^2(0, T)$ para $i = 1, \dots, n$. Si $u : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ es continua, con derivadas parciales $u_t, u_{x_i}, u_{x_i x_j}$ continuas para $i, j = 1, \dots, n$, entonces

$$d(u(t, X^1, \dots, X^n)) = u_t dt + \sum_{i=1}^n u_{x_i} dX^i + \frac{1}{2} \sum_{i,j=1}^n u_{x_i x_j} G^i G^j dt. \quad (3.27)$$

Ecuaciones diferenciales estocásticas

4.1. Definición y ejemplos

Sea $W(\cdot)$ un proceso de Wiener unidimensional y X_0 una variable aleatoria n -dimensional independiente de $W(\cdot)$. De ahora en adelante tomaremos $\mathcal{F}(t) := \mathcal{U}(W(s) \mid 0 \leq s \leq t, X_0), t \geq 0$, la σ -álgebra generada por X_0 y la historia del proceso de Wiener hasta el tiempo t . Consideramos además $T > 0$ y $b, B : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ con componentes $b = (b^1, b^2, \dots, b^n)^T$, $B = (B^1, B^2, \dots, B^n)^T$.

Definición 4.1. *Un proceso estocástico X en \mathbb{R}^n es solución de la ecuación diferencial estocástica (EDE) de Itô*

$$\begin{cases} dX = b(t, X)dt + B(t, X)dW \\ X(0) = X_0 \end{cases}, \quad 0 \leq t \leq T, \quad (4.1)$$

si $X(\cdot)$ es progresivamente medible con respecto a $\mathcal{F}(\cdot)$, $F := b(t, X) \in \mathbb{L}_n^1(0, T)$, $G := B(t, X) \in \mathbb{L}_n^2(0, T)$ y

$$X(t) = X_0 + \int_0^t b(s, X(s))ds + \int_0^t B(s, X(s))dW, \quad \forall t \in [0, T], \text{ casi seguro.} \quad (4.2)$$

Ejemplo 4.2. (Ecuación del movimiento Browniano geométrico). Si $S(t)$ denota el precio de una acción en tiempo t , podemos modelizar la evolución de $S(t)$ suponiendo que $\frac{dS}{S}$, el cambio relativo del precio, evoluciona como

$$\begin{cases} \frac{dS}{S} = \mu dt + \sigma dW \\ S(0) = s_0 \end{cases} \quad (4.3)$$

con $s_0, \mu, \sigma > 0$ constantes, denominadas deriva y volatilidad del precio, respectivamente. s_0 representa el precio inicial de la acción. A continuación utilizaremos la regla de la cadena de Itô para resolver esta ecuación diferencial estocástica. Sea $Y = u(t, S) = \log S$. Entonces

$$dY = u_s dS + \frac{1}{2} u_{ss} \cdot G^2 dt = \frac{1}{S} [\mu \cdot S dt + \sigma \cdot S dW] + \frac{1}{2} \left(-\frac{1}{S^2}\right) (\sigma^2 \cdot S^2) dt, \text{ esto es,}$$

$$\begin{cases} dY = (\mu - \frac{1}{2}\sigma^2) dt + \sigma dW \\ Y(0) = \log s_0. \end{cases}$$

De aquí, $\log S = \log s_0 + (\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)$, y $S(t) = s_0 \cdot e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)}$. Veamos ahora cuál es el valor esperado y varianza para $S(t)$. Dado que

$$S(t) = s_0 + \mu \int_0^t S ds + \sigma \int_0^t S dW,$$

sigue usando (3.15) que $E[S(t)] = s_0 + \mu \int_0^t E[S(s)] ds$. Definiendo $e(t) = E[S(t)]$, se tiene que $\begin{cases} e'(t) = \mu e(t) \\ e(0) = s_0 \end{cases}$. Luego, $E[S(t)] = s_0 e^{\mu t}$, $t \geq 0$. Veamos ahora cómo obtener $V[S(t)]$. Dado que

$$S(t) = s_0 e^{(\mu - \frac{\sigma^2}{2}) \cdot t + \sigma \cdot W_t} = e^{(\mu - \frac{\sigma^2}{2}) \cdot t + \sigma \cdot W_t + (\log s_0)}, \quad (4.4)$$

sea $X(t) = (\mu - \frac{\sigma^2}{2}) \cdot t + (\ln s_0) + \sigma \cdot W_t$. Es claro que $X \sim N((\mu - \frac{\sigma^2}{2}) \cdot t + \log s_0, \sigma^2 t)$. Usando la función generatriz de una distribución normal (1.10) se deduce que:

$$E[S] = E[e^X] = e^{(\mu - \frac{\sigma^2}{2}) \cdot t + (\ln s_0) + \frac{1}{2} \sigma^2 t} = s_0 e^{\mu t},$$

$$E[S^2] = E[e^{2X}] = e^{2((\mu - \frac{\sigma^2}{2}) \cdot t + (\ln s_0)) + 2\sigma^2 t} = s_0^2 e^{(2\mu + \sigma^2)t}.$$

Luego, $V[S] = E[S^2] - E[S]^2 = s_0^2 \cdot e^{2\mu t} [e^{\sigma^2 t} - 1]$.

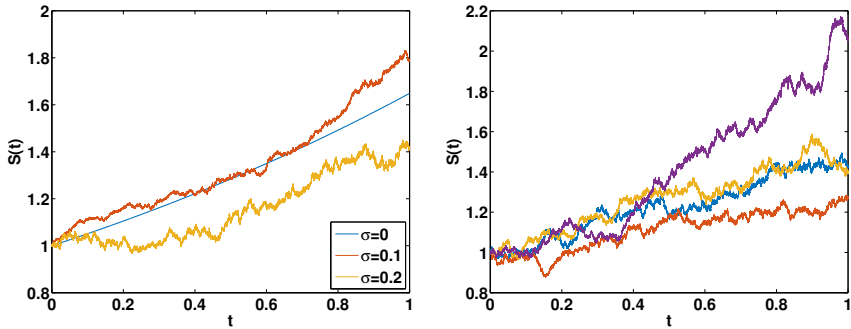


Figura 4.1. Una realización de la solución de la EDE (4.3) con $s_0 = 1$ para diferentes valores de σ (izquierda) y cuatro realizaciones de la solución de la EDE con $\mu = 0'5$, $\sigma = 0'2$ (derecha).

Ejemplo 4.3. (Puente Browniano). Consideramos la EDE

$$\begin{cases} dB &= -\frac{B}{1-t} dt + \sigma dW \\ B(0) &= 0 \end{cases} \quad (4.5)$$

y buscamos $B = X_1 \cdot X_2$ con

$$\begin{cases} dX_1 &= -\frac{1}{1-t} \cdot X_1 dt + 0 \cdot X_1 dW \\ X_1(0) &= 1 \end{cases} \quad \text{y} \quad \begin{cases} dX_2 &= f(t) dt + g(t) dW \\ X_2(0) &= B(0) = 0 \end{cases}$$

con f y g a determinar. Puesto que la primera ecuación es ordinaria, es claro que, $X_1(t) = 1 - t$. Ahora, por la fórmula del producto de Itô (3.22)

$$\begin{aligned} d(X_1 \cdot X_2) &= X_1 dX_2 + X_2 dX_1 + (0 \cdot X_1) g(t) dt \\ &= (1-t)[f(t) dt + g(t) dW] + X_2 \cdot \left[\frac{-1}{1-t} X_1 dt \right] \\ &= \left(\frac{-B}{1-t} + (1-t)f(t) \right) dt + (1-t)g(t) dW. \end{aligned}$$

Como $d(X_1 \cdot X_2) = dB$, tomando $f(t) = 0$ y $g(t) = \frac{\sigma}{1-t}$, sigue que $\begin{cases} dX_2 = \frac{\sigma}{1-t}dW \\ X_2(0) = 0 \end{cases}$ y $X_2(t) = \sigma \int_0^t \frac{1}{1-s}dW$. Luego, $B(t) = \sigma(1-t) \cdot \int_0^t \frac{1}{1-s}dW$, $0 \leq t < 1$, con $\lim_{t \rightarrow 1^-} B(t) = 0$, c.s.

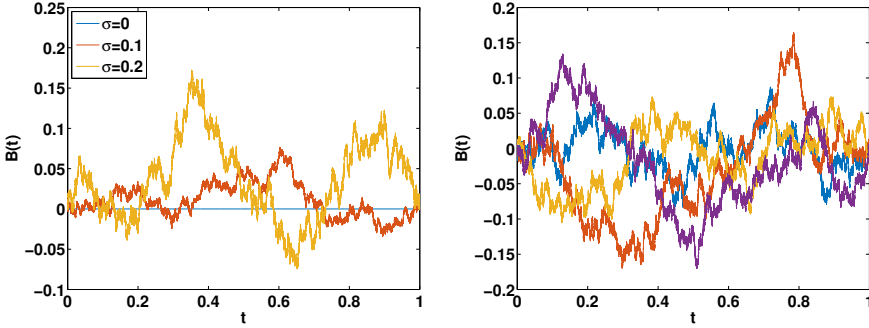


Figura 4.2. Una realización de la solución de la EDE (4.5) con $B_0 = 0$ para diferentes valores de σ (izquierda) y cuatro realizaciones de la solución de la EDE con $\sigma = 0.2$ (derecha).

Ejemplo 4.4. (Ecuación de Langevin). La velocidad $X(\cdot)$ de una partícula que sigue un movimiento Browniano con coeficiente de rozamiento $\mu > 0$, coeficiente de difusión σ y una velocidad inicial X_0 satisface

$$\begin{cases} dX = -\mu \cdot Xdt + \sigma dW \\ X(0) = X_0 \end{cases} \quad (4.6)$$

Para resolver (4.6), buscamos $X = X_1 \cdot X_2$ con

$$\begin{cases} dX_1 = -\mu X_1 dt \\ X_1(0) = 1 \end{cases} \quad \text{y} \quad \begin{cases} dX_2 = f(t)dt + g(t)dW \\ X_2(0) = X_0 \end{cases} \quad \text{con } f(t) \text{ y } g(t) \text{ a determinar.}$$

Luego $X_1(t) = e^{-\mu t}$ y

$$dX = X_1 dX_2 + X_2 dX_1 = (e^{-\mu t} \cdot f(t) - \mu X) dt + e^{-\mu t} \cdot g(t) dW.$$

Tomando $f \equiv 0$ y $g(t) = \sigma \cdot e^{\mu t}$ sigue que $X_2 = X_0 + \int_0^t \sigma e^{\mu s} dW$. Así

$$X(t) = e^{-\mu t} \cdot X_0 + \sigma \int_0^t e^{\mu(s-t)} dW, \quad t \geq 0.$$

Ejemplo 4.5. (Proceso de Ornstein Uhlenbek). Si $X(t)$ denota la posición de una partícula en tiempo t cuya velocidad es $V(t)$, entonces

$$\begin{cases} dX = V dt, & X(0) = X_0 \\ dV = -\mu V dt + \sigma dW, & V(0) = V_0 \end{cases} \quad (4.7)$$

donde X_0, V_0 son variables aleatorias y además $\mu > 0$ es el coeficiente de rozamiento y σ el coeficiente de difusión. Observamos que la velocidad de la partícula $V(t) =$

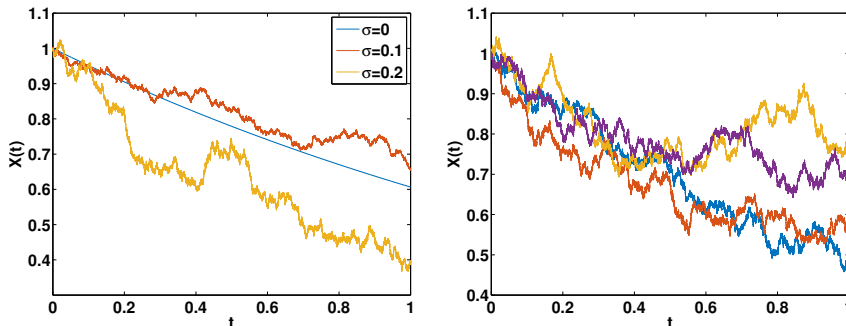


Figura 4.3. Una realización de la solución de la EDE (4.6) con $X_0 = 1$ para diferentes valores de σ (izquierda) y cuatro realizaciones de la solución de la EDE con $\mu = 0'5$, $\sigma = 0'2$ (derecha).

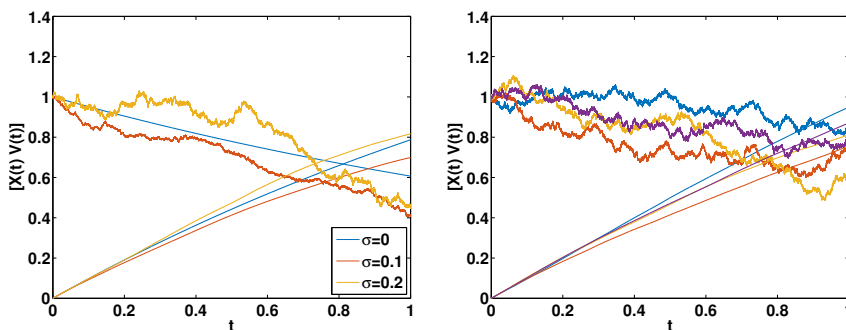


Figura 4.4. Una realización de la solución de la EDE (4.7) con $X_0 = 0$, $V_0 = 1$ para diferentes valores de σ (izquierda) y cuatro realizaciones de la solución de la EDE con $\mu = 0'5$, $\sigma = 0'2$ (derecha).

$e^{-\mu t} \cdot V_0 + \sigma \int_0^t e^{(s-t)} dW$, $t \geq 0$, satisface la ecuación de Langevin (4.6). Ahora, la posición de la partícula viene dada por $X(t) = X_0 + \int_0^t V(s) ds$. En particular,

$$E[X(t)] = E[X_0] + \int_0^t E[V(s)] ds = E[X_0] + \int_0^t e^{-\mu s} E[V_0] ds = E[X_0] + \frac{1 - e^{-\mu t}}{\mu} E[V_0].$$

Ejemplo 4.6. (Oscilador Armónico Estocástico). Consideramos la EDE

$$\begin{cases} dX = -V dt, & X(0) = X_0 \\ dV = (-\lambda^2 X - \mu V) dt + \sigma dW, & V(0) = V_0 \end{cases}, \quad (4.8)$$

donde $-\lambda^2 X$ representa una fuerza recuperadora y $-\mu V$ un término de rozamiento. En forma vectorial:

$$\begin{bmatrix} dX \\ dV \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ -\lambda^2 & \mu \end{bmatrix}}_D \begin{bmatrix} X \\ V \end{bmatrix} dt + \begin{bmatrix} 0 \\ \sigma \end{bmatrix} dW$$

de modo que (ver [3, p. 98-99])
$$\begin{bmatrix} X(t) \\ V(t) \end{bmatrix} = e^{Dt} \begin{bmatrix} X_0 \\ V_0 \end{bmatrix} + \int_0^t e^{D(t-s)} \begin{bmatrix} 0 \\ \sigma \end{bmatrix} dW.$$

Ahora cuando $\mu = 0$, $e^{Ds} = \begin{bmatrix} \cos(\lambda s) & \frac{1}{\lambda} \text{sen}(\lambda s) \\ -\lambda \text{sen}(\lambda s) & \cos(\lambda s) \end{bmatrix}$ y

$$\begin{bmatrix} X(t) \\ V(t) \end{bmatrix} = \begin{bmatrix} \cos(\lambda t) \cdot X_0 + \frac{1}{\lambda} \text{sen}(\lambda t) \cdot V_0 \\ -\lambda \text{sen}(\lambda t) \cdot X_0 + \cos(\lambda t) \cdot V_0 \end{bmatrix} + \sigma \int_0^t \begin{bmatrix} \frac{1}{\lambda} \text{sen}(\lambda(t-s)) \\ \cos(\lambda(t-s)) \end{bmatrix} dW.$$

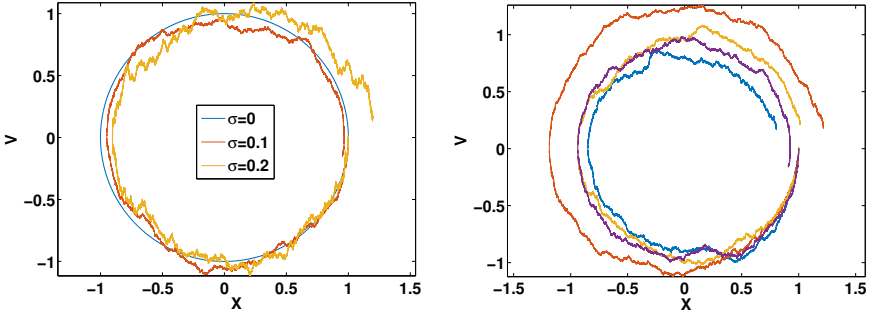


Figura 4.5. Una realización de la órbita de la EDE (4.8) con $X_0 = 1, V_0 = 0$ para diferentes valores de σ (izquierda), con $\mu = 0$ y $\lambda = 1$, y cuatro realizaciones de la órbita de la EDE con $\mu = 0, \lambda = 1, \sigma = 0'2$ (derecha).

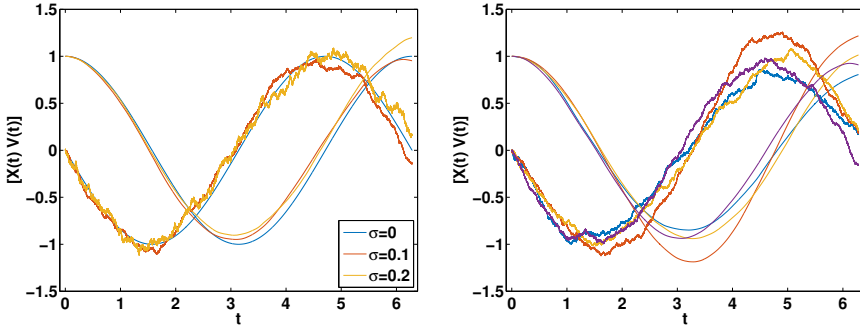


Figura 4.6. Una realización de la solución de la EDE (4.8) con $X_0 = 1, V_0 = 0$ para diferentes valores de σ (izquierda), con $\mu = 0$ y $\lambda = 1$ y cuatro realizaciones de la solución de la EDE con $\mu = 0, \lambda = 1, \sigma = 0'2$ (derecha).

4.2. Existencia y unicidad de solución

En esta sección estudiaremos la existencia y unicidad de solución para la EDE (4.1) en el sentido de (4.2). Para ello necesitaremos tener en cuenta el siguiente lema clásico.

Lema 4.7. (Desigualdad de Gronwall [3, p. 92]). Sean ϕ y f funciones continuas no negativas definidas en $0 \leq t \leq T$, y $C_0 \geq 0$ una constante. Si $\phi(t) \leq C_0 + \int_0^t f(s)\phi(s)ds$, para todo $0 \leq t \leq T$, entonces $\phi(t) \leq C_0 e^{\int_0^t f(s)ds}$ para todo $0 \leq t \leq T$.

Teorema 4.8. (*Existencia y unicidad de solución*). Sean b y $B : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ continuas cumpliendo las siguientes condiciones para todo $0 \leq t \leq T$, $x, \hat{x} \in \mathbb{R}^n$

$$\begin{aligned} a) & |b(t, x) - b(t, \hat{x})| \leq L|x - \hat{x}| \quad y \quad |B(t, x) - B(t, \hat{x})| \leq L|x - \hat{x}|, \\ b) & |b(t, x)| \leq L(1 + |x|) \quad y \quad |B(t, x)| \leq L(1 + |x|), \end{aligned}$$

para alguna constante L independiente de T . Sea además X_0 una variable aleatoria en \mathbb{R}^n verificando que

$$\begin{aligned} c) & E(|X_0|^2) < \infty \\ d) & X_0 \text{ es independiente de } \mathcal{W}^+(0) \end{aligned}$$

donde $W(\cdot)$ es un proceso de Wiener unidimensional. Entonces existe una única solución $X \in \mathbb{L}_n^2(0, T)$ de la EDE (4.1).

Observación 4.9. Unicidad de solución para (4.1) significa que si $X, \hat{X} \in \mathbb{L}_n^2(0, T)$ con trayectorias continuas casi seguro son soluciones de la EDE entonces, $P(X(t) = \hat{X}(t), \forall 0 \leq t \leq T) = 1$.

Demostración. i) Unicidad.

Supongamos que X y \hat{X} son soluciones de (4.1). Entonces para todo $0 \leq t \leq T$,

$$X(t) - \hat{X}(t) = \int_0^t (b(s, X) - b(s, \hat{X})) ds + \int_0^t (B(s, X) - B(s, \hat{X})) dW \quad (4.9)$$

A continuación veamos que $E[|X(t) - \hat{X}(t)|^2] \leq 2L^2 \cdot (T+1) \cdot \int_0^t E[|X(s) - \hat{X}(s)|^2] ds$. En efecto, como $(a+b)^2 \leq 2a^2 + 2b^2$ tenemos de (4.9) que

$$E[|X(t) - \hat{X}(t)|^2] \leq 2E \left[\left| \int_0^t (b(s, X) - b(s, \hat{X})) ds \right|^2 \right] + 2E \left[\left| \int_0^t (B(s, X) - B(s, \hat{X})) dW \right|^2 \right].$$

En primer lugar, usando la desigualdad de Cauchy-Schwartz y (a)

$$\begin{aligned} E \left[\left| \int_0^t (b(s, X) - b(s, \hat{X})) ds \right|^2 \right] &\leq TE \left[\int_0^t |b(s, X) - b(s, \hat{X})|^2 ds \right] \\ &\leq L^2 T \int_0^t E[|X(s) - \hat{X}(s)|^2] ds. \end{aligned}$$

Ahora, usando la propiedad de isometría de Itô (3.16) y (a)

$$\begin{aligned} E \left[\left| \int_0^t (B(s, X) - B(s, \hat{X})) dW \right|^2 \right] &= E \left[\int_0^t |B(s, X) - B(s, \hat{X})|^2 ds \right] \\ &\leq L^2 \int_0^t E[|X(s) - \hat{X}(s)|^2] ds. \end{aligned}$$

Por lo tanto, con $C = 2L^2(T+1)$, $E[|X(t) - \hat{X}(t)|^2] \leq C \int_0^t E[|X(s) - \hat{X}(s)|^2] ds$, $0 \leq t \leq T$. En virtud del Lema 4.7, $E[|X(t) - \hat{X}(t)|^2] = 0$, para todo $t \in [0, T]$.

Así $X(t) = \hat{X}(t)$ c.s., y como X y \hat{X} tienen trayectorias continuas casi seguro, $P(\max_{0 \leq t \leq T} |X(t) - \hat{X}(t)| > 0) = 0$.

ii) Existencia.

$$1) \text{ Definimos } X^0(t) := X_0, \quad X^{k+1}(t) := X_0 + \int_0^t b(s, X^k(s)) ds + \int_0^t B(s, X^k(s)) dW, \quad k \geq 0. \quad (4.10)$$

Sea $d^k(t) := E[|X^{k+1}(t) - X^k(t)|^2]$ y veamos que existe $M > 0$ tal que

$$d^k(t) \leq \frac{(M \cdot t)^{k+1}}{(k+1)!}, \quad \forall k \geq 0, \quad \forall t \in [0, T]. \quad (4.11)$$

$$\begin{aligned} \text{En primer lugar, } d^0(t) &= E[|X^1(t) - X^0(t)|^2] = E \left[\left| \int_0^t b(s, X_0) ds + \int_0^t B(s, X_0) dW \right|^2 \right] \leq \\ &\leq 2E \left[\left| \int_0^t b(s, X_0) ds \right|^2 \right] + 2E \left[\left| \int_0^t B(s, X_0) dW \right|^2 \right]. \end{aligned}$$

Usando nuevamente la desigualdad de Cauchy-Schwartz en la primera integral y la propiedad de isometría (3.16) en la segunda esperanza:

$$d^k(t) \leq 2E \left[t \cdot \int_0^t |b(s, X_0)|^2 ds \right] + 2E \left[\int_0^t |B(s, X_0)|^2 ds \right].$$

Usando (b), $d^k(t) \leq 2tE \left[\int_0^t L^2(1 + |X_0|)^2 ds \right] + 2E \left[\int_0^t L^2(1 + |X_0|)^2 ds \right] = Mt$, con $M := 2L^2(T+1)E[(1 + |X_0|)^2] \geq 2L^2(T+1)$. Observar que $E[(1 + |X_0|)^2] \leq 2(1 + E[|X_0|^2]) < \infty$.

Seguidamente, por inducción sobre k , y considerando un desarrollo análogo en el que se usa (a)

$$\begin{aligned} d^k(t) &= E[|X^{k+1}(t) - X^k(t)|^2] \\ &= E \left[\left| \int_0^t (b(s, X^k(s)) - b(s, X^{k-1}(s))) ds + \int_0^t (B(s, X^k(s)) - B(s, X^{k-1}(s))) dW \right|^2 \right] \\ &\leq 2L^2(T+1) \cdot E \left[\int_0^t |X^k(s) - X^{k-1}(s)|^2 ds \right] \\ &\leq M \int_0^t \frac{(M \cdot s)^k}{k!} ds = \frac{M^{k+1} \cdot t^{k+1}}{(k+1)!}. \end{aligned}$$

2) Veamos ahora que

$$E[\max_{0 \leq t \leq T} |X^{k+1}(t) - X^k(t)|^2] \leq C \cdot \frac{(M \cdot T)^k}{k!}, \quad \text{con } C := 2L^2T(T+4). \quad (4.12)$$

Tenemos que

$$\begin{aligned} |X^{k+1}(t) - X^k(t)|^2 &\leq 2 \left| \int_0^t (b(s, X^k(s)) - b(s, X^{k-1}(s))) ds \right|^2 + \\ &\quad + 2 \left| \int_0^t (B(s, X^k(s)) - B(s, X^{k-1}(s))) dW \right|^2. \end{aligned}$$

Ahora, usando la desigualdad de Cauchy-Schwartz y (a) en la primera integral:

$$\begin{aligned} |X^{k+1}(t) - X^k(t)|^2 &\leq 2TL^2 \int_0^T |X^k(s) - X^{k-1}(s)|^2 ds + \\ &\quad + 2 \left| \int_0^t (B(s, X^k(s)) - B(s, X^{k-1}(s))) dW \right|^2. \end{aligned}$$

Tomando máximos y esperanzas, y usando (4.11) y la desigualdad de Doob L^2 (1.12)

$$\begin{aligned} E \left[\max_{0 \leq t \leq T} |X^{k+1}(t) - X^k(t)|^2 \right] &\leq 2TL^2 \int_0^T E[|X^k(s) - X^{k-1}(s)|^2] ds \\ &+ 2E \left[\max_{0 \leq t \leq T} \left| \int_0^t (B(s, X^k(s)) - B(s, X^{k-1}(s))) dW \right|^2 \right] \\ &\leq 2TL^2 \int_0^T \frac{(M \cdot s)^k}{k!} ds + 2 \cdot 2^2 \cdot E \left[\left| \int_0^T (B(s, X^k(s)) - B(s, X^{k-1}(s))) dW \right|^2 \right]. \end{aligned}$$

Ahora por la Isometría de Itô (3.16) y (a)

$$\begin{aligned} |X^{k+1}(t) - X^k(t)|^2 &\leq (2TL^2) \cdot \left(T \cdot \frac{M^k \cdot T^k}{k!} \right) + 8L^2 \int_0^T E[|X^k(s) - X^{k-1}(s)|^2] ds \\ &\leq 2L^2(T+4) \cdot \left(T \cdot \frac{M^k \cdot T^k}{k!} \right) = C \frac{(M \cdot T)^k}{k!}, \quad C := 2L^2T(T+4). \end{aligned}$$

3) A continuación, sea $A_k = \{w \in \Omega / \max_{0 \leq t \leq T} |X^{k+1}(t) - X^k(t)| > \frac{1}{2^k}\}$, $k \geq 0$. Teniendo en cuenta la desigualdad de Chebyshev (1.3) con $p = 2$ y (4.12)

$$P(A_k) \leq 2^{2k} E \left[\max_{0 \leq t \leq T} |X^{k+1}(t) - X^k(t)|^2 \right] \leq 2^{2k} \frac{C(MT)^k}{k!},$$

con $\sum_{k=0}^{\infty} 2^{2k} \frac{(MT)^k}{k!} < \infty$. Según el Lema 1.10, $P(\overline{\lim}_{k \rightarrow \infty} A_k) = 0$. Así, para casi todo ω ,

$X^k = X^0 + \sum_{j=0}^{k-1} (X^{j+1} - X^j)$ converge uniformemente en $[0, T]$ a un proceso estocástico

X . En la definición de X^{k+1} (4.10) pasamos al límite para probar que X cumple

$$X(t) = X_0 + \int_0^t b(s, X(s)) ds + \int_0^t B(s, X(s)) dW, \quad 0 \leq t \leq T, \quad \text{c.s.}$$

Además, como X^k tiene trayectorias continuas c.s. dicha propiedad también se cumple para X .

4) Veamos finalmente que $X \in \mathbb{L}_n^2(0, T)$, esto es $E \left[\int_0^T |X(t)|^2 dt \right] < \infty$. Usando que $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$, y la desigualdad de Cauchy-Schwartz, de (4.10):

$$\begin{aligned} |X^{k+1}(t)|^2 &\leq 3|X_0|^2 + 3 \left| \int_0^t b(s, X_k(s)) ds \right|^2 + 3 \left| \int_0^t B(s, X_k(s)) dW \right|^2 \\ &\leq 3|X_0|^2 + 3T \int_0^t |b(s, X_k(s))|^2 ds + 3 \left| \int_0^t B(s, X_k(s)) dW \right|^2 \\ &\leq 3|X_0|^2 + 6L^2T^2 + 6L^2T \int_0^t |X_k(s)|^2 ds + 3 \left| \int_0^t B(s, X_k(s)) dW \right|^2, \end{aligned}$$

donde en la última desigualdad hemos usado que, por (b), $|b(s, X^k(s))|^2 \leq L^2(1 + |X^k(s)|)^2 \leq 2L^2(1 + |X^k(s)|^2)$. Tomando esperanzas y aplicando nuevamente (b) junto con la isometría de Itô (3.16)

$$E[|X^{k+1}(t)|^2] \leq C(1 + E[|X_0|^2]) + C \int_0^t E[|X^k(s)|^2] ds, \quad \forall k \geq 0, \forall 0 \leq t \leq T, \quad (4.13)$$

siendo $C > 1$ constante que sólo depende de L y T . En particular,

$$\begin{aligned} E[|X^1(t)|^2] &\leq C \cdot (1 + E[|X_0|^2]) + C \cdot E[|X_0|^2] \cdot t \leq C \cdot (1 + E[|X_0|^2]) + C^2 \cdot (1 + E[|X_0|^2]) \cdot t \\ &\leq C \cdot (1 + E[|X_0|^2]) \cdot (1 + Ct) \leq C \cdot (1 + E[|X_0|^2])e^{C \cdot t}, \quad \forall t \in [0, T]. \end{aligned}$$

Inductivamente, si $E[|X^k(t)|^2] \leq C \cdot (1 + E[|X_0|^2])e^{C \cdot t}$, $0 \leq t \leq T$, entonces de (4.13)

$$E[|X^{k+1}(t)|^2] \leq C \cdot (1 + E[|X_0|^2]) \cdot \left\{ 1 + C \int_0^t e^{C \cdot s} ds \right\} = C(1 + E[|X_0|^2])e^{C \cdot t}, \quad 0 \leq t \leq T.$$

Haciendo $k \rightarrow \infty$:

$$E[|X(t)|^2] \leq C(1 + E[|X_0|^2])e^{C \cdot t}, \quad t \in [0, T], \quad (4.14)$$

y de aquí que $E \left[\int_0^T |X(t)|^2 dt \right] < \infty$. □

Observación 4.10. Respecto a la solución de (4.1), se verifica además que

$$E[|X(t) - X(s)|^2] \leq C \cdot (t - s), \quad 0 \leq s \leq t \leq T, \quad (4.15)$$

para cierta constante C dependiente de L , T y $E[X_0^2]$. En efecto, observemos que $X(t) - X(s) = \int_s^t b(u, X(u))du + \int_s^t B(u, X(u))dW$. Luego, usando la desigualdad de Cauchy-Schwartz, la propiedad de isometría de Itô (3.16), (b) y (4.14):

$$\begin{aligned} E[|X(t) - X(s)|^2] &\leq 2E \left[\left(\int_s^t b(u, X(u))du \right)^2 \right] + 2E \left[\left(\int_s^t B(u, X(u))dW \right)^2 \right] \\ &\leq C \left\{ E \left[\int_s^t |b(u, X(u))|^2 du \right] + E \left[\int_s^t |B(u, X(u))|^2 du \right] \right\} \\ &\leq C \cdot E \left[\int_s^t (1 + |X(u)|^2) du \right] = C \cdot \left\{ (t - s) + \int_s^t E[|X(u)|^2] du \right\} \\ &\leq C \cdot \left\{ (t - s) + \int_s^t e^{C \cdot u} du \right\} \leq C \cdot \left\{ (t - s) + e^{C \cdot T} (t - s) \right\} \\ &\leq C \cdot (t - s), \quad 0 \leq s \leq t \leq T, \end{aligned}$$

donde la constante C se ha tomado acumulativa, pero es finita y sólo depende de L , T y $E[X_0]$. Finalizamos esta sección recogiendo algunas propiedades de la solución de (4.1).

Teorema 4.11. (*Acotación de los momentos*). *En las mismas condiciones del Teorema 4.8, si además $E[X_0^{2p}] < \infty$ para algún $p > 1$ natural, entonces la solución X verifica*

$$E[|X(t)|^{2p}] \leq C_2(1 + E[X_0^{2p}])e^{C_1 t}, \quad E[(X(t) - X_0)^{2p}] \leq C_2(1 + E[X_0^{2p}])t^p e^{C_1 t},$$

para $0 \leq t \leq T$ y ciertas constantes $C_i = C_i(L, T, n, p)$, $i = 1, 2$.

Aunque no incluimos la prueba (véase [8, p. 136-138] para más detalles), esta está basada en aplicar la fórmula de Itô (3.26) a $(X_t)^{2p}$ para obtener

$$\begin{aligned} (X(t))^{2p} &= (X_0)^{2p} + \int_0^t [2p(X(s))^{2p-1}b(s, X(s)) + p(2p-1)(X(s))^{2p-2}B(s, X(s))^2]ds \\ &\quad + \int_0^t 2p(X(s))^{2p-1}B(s, X(s))dW. \end{aligned}$$

4.3. Métodos numéricos elementales para EDEs

4.3.1. El método de Euler-Maruyama

Consideramos la ecuación estocástica de Itô (4.1), donde X_0 , b y B satisfacen al menos la hipótesis del Teorema 4.8, y W es un proceso de Wiener unidimensional. Consideramos una partición del intervalo $[0, T]$

$$0 = t_0 < t_1 < \dots < t_N = T, \quad \text{con } h := \max_{0 \leq j \leq N-1} (t_{j+1} - t_j). \quad (4.16)$$

En dicha partición queremos construir un proceso estocástico discreto $X^h = \{X_j^h\}_{j=0}^N$ de tal modo que $X_j^h \simeq X_{t_j}$, $0 \leq j \leq N$, en algún sentido que precisaremos a continuación. A lo largo de esta sección, para simplificar la notación, denotaremos por X_t a la variable aleatoria que corresponde al proceso estocástico en tiempo t ($X(t)$).

Definición 4.12. *Dada la partición (4.16) de $[0, T]$, un proceso estocástico discreto $X^h = \{X_j^h\}_{j=0}^N$ converge fuertemente a X solución de (4.1) si*

$$\lim_{h \rightarrow 0} \max_{0 \leq j \leq N} E[|X_j^h - X_{t_j}|] = 0.$$

Además, la convergencia fuerte es de orden γ si

$$\max_{0 \leq j \leq N} E[|X_j^h - X_{t_j}|] \leq C \cdot h^\gamma, \quad 0 < h < h_0,$$

siendo C constante positiva independiente de h .

Definición 4.13. *Bajo las mismas condiciones de la definición previa, X^h converge débilmente a X si*

$$\lim_{h \rightarrow 0} \max_{0 \leq j \leq N} |E[g(X_j^h)] - E[g(X_{t_j})]| = 0, \quad \text{para toda función } g \text{ Lipschitziana.}$$

Además, la convergencia débil es de orden β si

$$\max_{0 \leq j \leq N} |E[g(X_j^h)] - E[g(X_{t_j})]| \leq C_g \cdot h^\beta, \quad 0 < h < h_0,$$

donde C_g es una constante positiva independiente de h .

Observación 4.14. La definición anterior se podría generalizar considerando funciones g con crecimiento polinómico (véase [8]).

Teorema 4.15. *Si X^h converge a X fuertemente con orden γ y débilmente con orden β , entonces $\beta \geq \gamma$.*

Demostración. Observamos que $|E[g(X_j^h)] - E[g(X_{t_j})]| \leq E[|g(X_j^h) - g(X_{t_j})|]$. Si L_g es una constante de Lipschitz de g $|g(X_j^h) - g(X_{t_j})| \leq L_g \cdot |X_j^h - X_{t_j}|$. Luego, $|E[g(X_j^h)] - E[g(X_{t_j})]| \leq L_g \cdot E[|X_j^h - X_{t_j}|]$, de lo cual se deduce inmediatamente el resultado. \square

Observación 4.16. Observar que la convergencia fuerte garantiza la proximidad de las trayectorias para una misma realización de los procesos X^h y X , mientras que la convergencia débil solo asegura la proximidad de las trayectorias $E[g(X^h)]$ y $E[g(X)]$ (para toda g Lipschitziana).

Definición 4.17. Se llama método de Euler-Maruyama al método numérico por el cual se genera en la partición (4.16) el proceso discreto

$$\begin{cases} X_{n+1}^h = X_n^h + b(t_n, X_n^h) \cdot h_n + B(t_n, X_n^h) \cdot \Delta W_n \\ X_0^h \text{ dado,} \end{cases}$$

siendo $\Delta W_n := W(t_{n+1}) - W(t_n) \sim N(0, h_n)$ variables aleatorias independientes, y $h_n = t_{n+1} - t_n, 0 \leq n \leq N - 1$.

Observación 4.18. Claramente, el método de Euler-Maruyama surge de considerar

$$X_{t_{n+1}} = X_{t_n} + \int_{t_n}^{t_{n+1}} b(s, X_s) ds + \int_{t_n}^{t_{n+1}} B(s, X_s) dW_s$$

y aproximar las integrales mediante $\int_{t_n}^{t_{n+1}} b(s, X_s) ds \simeq b(t_n, X_{t_n}) \cdot h_n$ y $\int_{t_n}^{t_{n+1}} B(s, X_s) dW_s \simeq B(t_n, X_{t_n}) \cdot \int_{t_n}^{t_{n+1}} dW_s = B(t_n, X_{t_n}) \cdot \Delta W_n$.

Observación 4.19. El siguiente resultado nos dice que el método de Euler-Maruyama converge a la solución de (4.1) con orden fuerte (al menos) $\gamma = \frac{1}{2}$. En particular, su orden de convergencia débil será $\beta \geq \frac{1}{2}$. De hecho el orden débil de dicho método es $\beta = 1$, aunque probar dicha propiedad se escapa de los objetivos de este trabajo. No obstante, ilustraremos estos órdenes de convergencia al final de éste capítulo.

Teorema 4.20. En las mismas condiciones del Teorema (4.8), si además b y B verifican

$$|b(t, x) - b(s, y)|^2 \leq L \cdot (|x - y|^2 + |t - s|), \quad |B(t, x) - B(s, y)|^2 \leq L \cdot (|x - y|^2 + |t - s|) \tag{4.17}$$

donde $0 \leq s, t \leq T, x, y \in \mathbb{R}^n$, entonces el método de Euler-Maruyama converge fuertemente con orden al menos $\frac{1}{2}$, esto es, si $E[|X_0^h - X_0|^2]^{\frac{1}{2}} \leq K \cdot h^{\frac{1}{2}}$, para alguna constante K independiente de h , entonces

$$\max_{0 \leq j \leq N} E[|X_j^h - X_{t_j}|] \leq C \cdot h^{\frac{1}{2}},$$

donde C solo depende de L, T, K y $E[X_0^2]$.

Demostración. Sean $\{X_k^h\}_{k=0}^N$ las iteraciones del método de Euler-Maruyama aplicado sobre la partición (4.16), y consideramos el proceso estocástico continuo definido a trozos por

$$X_t^h = X_k^h + b(t_k, X_k^h) \cdot (t - t_k) + B(t_k, X_k^h) \cdot (W_t - W_{t_k}), \quad t \in [t_k, t_{k+1}), \quad 0 \leq k \leq N - 1,$$

con X_0^h tal que $E[|X_0^h - X_0|^2] \leq K^2 \cdot h$. Sean b^h y B^h definidas como

$$b^h(s, Y) = b(t_k, Y_{t_k}), \quad B^h(s, Y) = B(t_k, Y_{t_k}), \quad \text{para } s \in [t_k, t_{k+1}), \quad 0 \leq k \leq N - 1.$$

Entonces:

$$X_s^h = X_0^h + \int_0^s b^h(u, X_u^h) du + \int_0^s B^h(u, X_u^h) dW_u, \quad \forall s \in [0, T],$$

y, por tanto

$$X_s^h - X_s = (X_0^h - X_0) + \int_0^s [b^h(u, X_u^h) - b(u, X_u)] du + \int_0^s [B^h(u, X_u^h) - B(u, X_u)] dW_u.$$

Usando que $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$ y la desigualdad de Cauchy-Schwartz para la integral de Riemann sigue que:

$$(X_s^h - X_s)^2 \leq 3(X_0^h - X_0)^2 + 3s \int_0^s [b^h(u, X_u^h) - b(u, X_u)]^2 du + 3 \left(\int_0^s [B^h(u, X_u^h) - B(u, X_u)] dW_u \right)^2. \quad (4.18)$$

Ahora, dado $u \in [0, s]$, existe $k \in \{0, \dots, N - 1\}$ tal que $u \in [t_k, t_{k+1})$, y entonces:

$$\begin{aligned} |b^h(u, X_u^h) - b(u, X_u)|^2 &= |b(t_k, X_{t_k}^h) - b(u, X_u)|^2 \\ &\leq L|X_k^h - X_u|^2 + L(u - t_k) \leq \\ &\leq 2L|X_k^h - X_{t_k}|^2 + 2L|X_{t_k} - X_u|^2 + Lh. \end{aligned}$$

Sea $\varphi^h(t) := \sup_{0 \leq s \leq t} E[|X_s^h - X_s|^2]$, $t \in [0, T]$. Así, usando (4.15)

$$\begin{aligned} E[|b^h(u, X_u^h) - b(u, X_u)|^2] &\leq 2LE[|X_k^h - X_{t_k}|^2] + 2LE[|X_{t_k} - X_u|^2] + Lh \\ &\leq 2L\varphi^h(t_k) + 2LC(u - t_k) + Lh \\ &\leq 2L\varphi^h(u) + L(1 + 2C)h. \end{aligned}$$

Igualmente, $E[|B^h(u, X_u^h) - B(u, X_u)|^2] \leq 2L\varphi^h(u) + L(1 + 2C)h$. Luego, tomando esperanzas en (4.18) y usando la propiedad de isometría para la integral de Itô (3.16)

$$\begin{aligned} E[|X_s^h - X_s|^2] &\leq 3K^2h + 3s \int_0^s [L(1 + 2C)h + 2L\varphi^h(u)] du \\ &\quad + 3 \int_0^s [L(1 + 2C)h + 2L\varphi^h(u)] du \\ &\leq 3K^2h + 3(T + 1)TL(1 + 2C)h + 6(T + 1)L \int_0^s \varphi^h(u) du \\ &= C_0h + \int_0^s C_1\varphi^h(u) du, \end{aligned}$$

para todo $s \in [0, T]$, con $C_0 = 3K^2 + 3(T + 1)TL(1 + 2C)$ y $C_1 = 6(T + 1)L$. Por tanto:

$$\varphi^h(t) = \sup_{0 \leq s \leq t} E[|X_s^h - X_s|^2] \leq C_0 \cdot h + \int_0^t C_1\varphi^h(u) du, \quad \forall t \in [0, T],$$

y en virtud del Lema 4.7

$$\varphi^h(t) \leq C_0 h e^{\int_0^t C_1 du} \leq C_0 e^{C_1 T} h, \quad \forall t \in [0, T].$$

Luego, $E[|X_k^h - X_{t_k}|^2] \leq C_0 e^{C_1 T} h$, $0 \leq k \leq N$, y por la desigualdad de Cauchy-Schwartz (1.2)

$$E[|X_k^h - X_{t_k}|] \leq E[|X_k^h - X_{t_k}|^2]^{\frac{1}{2}} \leq C \cdot h^{\frac{1}{2}}, \quad 0 \leq k \leq N,$$

donde C es constante independiente de h . □

Observación 4.21. En las mismas condiciones del teorema previo se prueba en [8, p.342-344] que

$$E[\max_{0 \leq k \leq N} |X_k^h - X_{t_k}|^2] \leq C \cdot h,$$

con C independiente de h . Así, dado $0 < \varepsilon < \frac{1}{2}$, por la desigualdad de Chebyshev (1.3)

$$P(\max_{0 \leq k \leq N} |X_k^h - X_{t_k}| \geq h^{\frac{1}{2}-\varepsilon}) \leq \frac{1}{h^{1-2\varepsilon}} E[\max_{0 \leq k \leq N} |X_k^h - X_{t_k}|^2] \leq C \cdot h^{2\varepsilon}.$$

Luego, para todo ε , $0 < \varepsilon < \frac{1}{2}$, $\lim_{h \rightarrow 0} P(\max_{0 \leq k \leq N} |X_k^h - X_{t_k}| \geq h^{\frac{1}{2}-\varepsilon}) = 0$.

De hecho, en [5] y [4] se demuestra que para todo ε , $0 < \varepsilon < \frac{1}{2}$,

$$\max_{0 \leq k \leq N} |X_k^h(w) - X_{t_k}(w)| \leq C_{\varepsilon, w} \cdot h^{\frac{1}{2}-\varepsilon}, \quad \text{para casi todo } w \in \Omega.$$

Así, aparte de la convergencia en media cuadrática, las realizaciones del método de Euler-Maruyama convergen, casi seguro, a las correspondientes realizaciones de la solución de la EDE (4.1), con orden de convergencia casi $\frac{1}{2}$.

4.3.2. El método de Milstein

Puesto que el orden de convergencia fuerte del método de Euler-Maruyama es demasiado pequeño, interesa obtener métodos numéricos de mayor orden fuerte. Esto se puede conseguir, como veremos a continuación, por ejemplo mediante la aplicación de la regla de la cadena de Itô. En efecto, sea X solución de la EDE (4.1)

$$X_t = X_{t_0} + \int_{t_0}^t b(s, X_s) ds + \int_{t_0}^t B(s, X_s) dW_s. \tag{4.19}$$

Entonces, si $f(t, x)$ es suficientemente regular, la fórmula de Itô (3.27) indica que

$$f(t, X_t) = f(t_0, X_{t_0}) + \int_{t_0}^t L^{(0)} f(s, X_s) ds + \int_{t_0}^t L^{(1)} f(s, X_s) dW_s, \tag{4.20}$$

donde los operadores $L^{(1)}$ y $L^{(0)}$ vienen definidos como (con $b = (b^1, \dots, b^n)^T$, $B = (B^1, \dots, B^n)^T$)

$$L^{(0)} := \frac{\partial}{\partial t} + \sum_{i=1}^n b^i \cdot \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n B^i B^j \frac{\partial^2}{\partial x_i \partial x_j} \quad \text{y} \quad L^{(1)} := \sum_{i=1}^n B^i \cdot \frac{\partial}{\partial x_i}.$$

Aplicando (4.20) a $b(s, X_s)$ y $B(s, X_s)$ en (4.19) sigue que

$$\begin{aligned} X_t = X_{t_0} + \int_{t_0}^t \left[b(t_0, X_{t_0}) + \int_{t_0}^s L^{(0)} b(u, X_u) du + \int_{t_0}^s L^{(1)} b(u, X_u) dW_u \right] ds \\ + \int_{t_0}^t \left[B(t_0, X_{t_0}) + \int_{t_0}^s L^{(0)} B(u, X_u) du + \int_{t_0}^s L^{(1)} B(u, X_u) dW_u \right] dW_s. \end{aligned}$$

$$X_t = X_{t_0} + b(t_0, X_{t_0})(t - t_0) + B(t_0, X_{t_0})(W(t) - W(t_0)) + R_1(t, t_0), \tag{4.21}$$

con resto

$$\begin{aligned} R_1(t, t_0) := \int_{t_0}^t \int_{t_0}^s L^{(0)} b(u, X_u) dud s + \int_{t_0}^t \int_{t_0}^s L^{(1)} b(u, X_u) dW_u ds + \\ + \int_{t_0}^t \int_{t_0}^s L^{(0)} B(u, X_u) dud W_s + \int_{t_0}^t \int_{t_0}^s L^{(1)} B(u, X_u) dW_u dW_s. \end{aligned}$$

Despreciando $R_1(t, t_0)$ se obtiene la formulación del método de Euler-Maruyama. Ahora se pueden obtener métodos de Itô-Taylor de más alto orden aplicando sucesivamente

la fórmula de Itô a las integrales en $R_1(t, t_0)$. Así, aplicando nuevamente la fórmula de Itô (4.20) a $L^{(1)}B(u, X_u)$ en $R_1(t, t_0)$ se obtiene que

$$X_t = X_{t_0} + b(t_0, X_{t_0})(t - t_0) + B(t_0, X_{t_0})(W(t) - W(t_0)) \\ + L^{(1)}B(t_0, X_{t_0}) \int_{t_0}^t \int_{t_0}^s dW_u dW_s + R_2(t, t_0),$$

donde $L^{(1)}B = \sum_{i=1}^n B^i \frac{\partial B}{\partial x_i} = B' \cdot B$, siendo B' la matriz jacobiana de B , y puesto que W es un proceso de Wiener unidimensional, usando (3.5),

$$\int_{t_0}^t \int_{t_0}^s dW_u dW_s = \int_{t_0}^t (W(s) - W(t_0)) dW_s = \frac{(W_t - W_{t_0})^2 - (t - t_0)}{2}.$$

Además

$$R_2(t, t_0) = \int_{t_0}^t \int_{t_0}^s L^{(0)}b(u, X_u) dud s + \int_{t_0}^t \int_{t_0}^s L^{(1)}b(u, X_u) dW_u d s + \int_{t_0}^t \int_{t_0}^s L^{(0)}B(u, X_u) dud W_s \\ + \int_{t_0}^t \int_{t_0}^s \int_{t_0}^u L^{(1)}L^{(0)}B(v, X_v) dv dW_u dW_s + \int_{t_0}^t \int_{t_0}^s \int_{t_0}^u L^{(1)}L^{(1)}B(v, X_v) dW_v dW_u dW_s.$$

Suprimiendo $R_2(t, t_0)$ se obtiene una aproximación de Itô-Taylor que posee orden de convergencia fuerte $\gamma = 1$ y orden débil $\beta = 1$ (véase [8]).

Definición 4.22. Se llama método de Milstein al método numérico por el cual se genera en la partición (4.16) el proceso discreto

$$\begin{cases} X_{n+1}^h = X_n^h + b(t_n, X_n^h)h_n + B(t_n, X_n^h)\Delta W_n + \frac{1}{2}(B' \cdot B)(t_n, X_n^h)((\Delta W_n)^2 - h_n) \\ X_0^h \text{ dado,} \end{cases}$$

siendo $\Delta W_n = W(t_{n+1}) - W(t_n) \sim N(0, h_n)$ variables aleatorias independientes, y $h_n = t_{n+1} - t_n, 0 \leq n \leq N - 1$.

Observación 4.23. Para las iteraciones del método de Milstein $\{X_j^h\}_{j=0}^N$, bajo las mismas hipótesis del Teorema 4.20, se tiene que si $E[|X_0^h - X_0|^2]^{\frac{1}{2}} \leq K \cdot h$, para alguna constante K independiente de h , entonces $\max_{0 \leq j \leq N} E[|X_j^h - X_{t_j}|] \leq C \cdot h$, con C independiente de h , esto es, el método de Milstein converge fuertemente con orden $\gamma = 1$. Además su orden de convergencia débil es precisamente $\beta = 1$, al igual que ocurre con el método de Euler-Maruyama. Una demostración de estas propiedades de convergencia puede verse en [8, cap 10]. Asimismo, respecto a las realizaciones del método de Milstein, éstas convergen casi seguro a las correspondientes realizaciones de la solución de la EDE (4.1) con orden de convergencia casi 1, esto es, para todo $\varepsilon, 0 < \varepsilon < 1$, ([7, Teorema 3.3])

$$\max_{0 \leq j \leq N} |X_j^h(w) - X_{t_j}(w)| \leq C_{\varepsilon, w} \cdot h^{1-\varepsilon}, \quad \text{para casi todo } w \in \Omega.$$

Observación 4.24. Si B no depende de x entonces $B' = 0$ y los métodos de Milstein y de Euler-Maruyama coinciden. En tal caso, este último método resultaría tener orden de convergencia fuerte $\gamma = 1$ sobre tales problemas.

4.3.3. Ilustración numérica

En esta sección ilustramos los órdenes de convergencia fuerte y débil de los métodos de Euler-Maruyama y de Milstein en la EDE (4.3)

$$\begin{cases} dS = \mu S dt + \sigma S dW \\ S(0) = 1, \end{cases}$$

con $t \in [0, 1]$, donde elegimos $\mu = 0'5$ y $\sigma = 0'2$. Para cada $j = 1, 2, \dots, 9$, consideramos una partición uniforme del intervalo $[0, 1]$ en $N = 100 \cdot 2^j$ subintervalos, esto es, $t_n = n \cdot h$, $0 \leq n \leq N$, con $h = \frac{1}{N}$. Para cada uno de los 9 valores de h aplicamos los métodos de Euler-Maruyama y Milstein para obtener aproximaciones $\{S_n^h(w)\}_{n=0}^N$ a la solución exacta $\{S_{t_n}(w)\}_{n=0}^N$ en los correspondientes nodos de la partición, donde la solución exacta de la EDE (4.1) viene dada por

$$S_t(w) = e^{0'2W_t(w)+0'48t}, \quad t \in [0, 1].$$

Pretendemos ilustrar los órdenes de convergencia en el tiempo final $T = 1$. Para ello, para cada una de las particiones temporales, consideramos $M = 25000$ realizaciones del proceso de Wiener, y estimamos (ver [6])

$$\begin{aligned} E[|S_N^h - S_N|] &\simeq \frac{1}{M} \cdot \sum_{i=1}^M |S_N^h(w_i) - S_N(w_i)| := e_f(h), \\ |E[S_N^h] - E[S_N]| &\simeq \left| \frac{1}{M} \cdot \sum_{i=1}^M S_N^h(w_i) - \frac{1}{M} \cdot \sum_{i=1}^M S_N(w_i) \right| := e_d(h), \end{aligned}$$

donde hemos denotado $S_N = S_{t_N} = S_1 = e^{0'2W_1(w)+0'48}$. En la tabla siguiente se muestran los valores obtenidos para $e_f(h)$ y $e_d(h)$ con los métodos de Euler-Maruyama ($e_{f,EM}(h), e_{d,EM}(h)$) y de Milstein ($e_{f,Mils}(h), e_{d,Mils}(h)$) para cada $h = \frac{1}{N}$, con $N = 100 \cdot 2^j$, $j = 1, 2, \dots, 9$. En la Figura 4.7 (izquierda) observamos que $e_f(h)$ se comporta como $h^{\frac{1}{2}}$ y como h^1 , para los métodos de Euler-Maruyama y de Milstein, respectivamente, para $h \rightarrow 0$. En la Figura 4.7 (derecha) se observa que para ambos métodos $e_d(h)$ se comporta como h^1 , $h \rightarrow 0$. Estas gráficas ilustran los ordenes de convergencia fuerte y débil de ambos métodos respectivamente.

h^{-1}	$e_{d,EM}(h)$	$e_{d,Mils}(h)$	$e_{f,EM}(h)$	$e_{f,Mils}(h)$
200	1.0441e-03	1.0313e-03	2.8134e-03	1.1139e-03
400	5.4229e-04	5.0996e-04	1.9302e-03	5.5211e-04
800	2.7367e-04	2.5493e-04	1.3388e-03	2.7612e-04
1600	1.2771e-04	1.2831e-04	9.3245e-04	1.3918e-04
3200	6.0930e-05	6.3599e-05	6.5582e-04	6.8975e-05
6400	3.7612e-05	3.2090e-05	4.6639e-04	3.4660e-05
12800	1.3678e-05	1.5987e-05	3.2838e-04	1.7302e-05
25600	7.6752e-06	7.9895e-06	2.3333e-04	8.6658e-06
51200	3.5820e-06	3.9706e-06	1.6373e-04	4.3020e-06

Tabla 4.1. Estimaciones de los errores débil ($e_d(h)$) y fuerte ($e_f(h)$), con $h = \frac{1}{N}$ y $N = 100 \cdot 2^j$, $1 \leq j \leq 9$, tomando para cada h 25000 realizaciones del proceso de Wiener W , en la EDE (4.3) con $\mu = 0'5$, $\sigma = 0'2$, $S(0) = 1$ y $T = 1$.

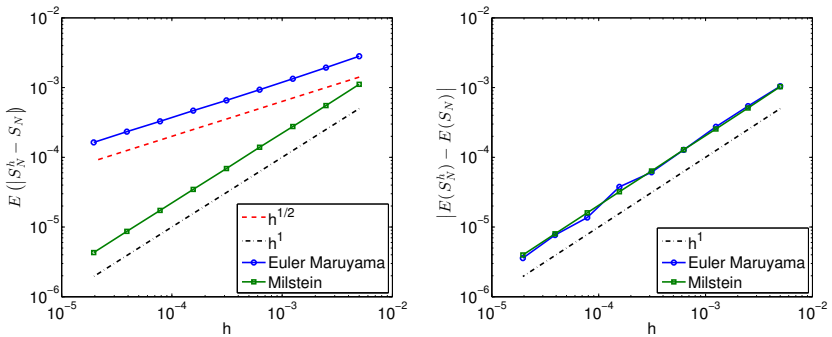


Figura 4.7. Estimaciones numéricas del orden fuerte (izquierda) y débil (derecha) de los métodos de Euler-Maruyama y Milstein en la EDE (4.3) con $\mu = 0'5$, $\sigma = 0'2$, $S(0) = 1$ y $T = 1$.

Bibliografia

- [1] L. Breiman, *Probability*, SIAM 1992.
- [2] Z. Brzezniak y T. Zastawniak, *Basic stochastic processes. A course through exercises*, Springer 1999.
- [3] L.C. Evans, *An Introduction to Stochastic Differential Equations*, AMS 2013.
- [4] G. Fleury, *Convergence of schemes for stochastic differential equations*, Prob. Engineering Mech. 21 (2005), p. 35-43.
- [5] I. Gyöngy, *A note on Euler's approximations*, Potential Anal. 8 (1998), p. 205-216.
- [6] D.J. Higham, *An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations*, SIAM Review 43-3 (2011), p. 525-546.
- [7] A. Jentzen y P.E. Kloeden, *Taylor Approximation for Stochastic Partial Differential Equations*, SIAM 2011.
- [8] P.E. Kloeden y E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer 1999.
- [9] A.N. Kolmogorov y S.V. Fomin, *Introductory Real Analysis*, Dover Publications, 1975.
- [10] B. Øksendal, *Stochastic Differential Equations. An Introduction with Applications*, Springer 1998.
- [11] Open Course Ware, *Advanced Stochastic Analysis. Lecture Notes*, Massachusetts Institute of Technology 2013.
<https://ocw.mit.edu/courses/sloan-school-of-management/15-070j-advanced-stochastic-processes-fall-2013/lecture-notes/>
- [12] W. Rudin, *Real and Complex Analysis*, McGraw-Hill 1987.

Introduction to Stochastic Differential

Equations and Numerical Methods



Sección de Matemáticas
Universidad de La Laguna

Raúl Melián Moreno

Facultad de Ciencias · Sección de Matemáticas

Universidad de La Laguna

alu0100989908@ull.edu.es

Abstract

In the present manuscript, we study stochastic differential equations. Initially, we will give a short introduction to the probability theory that is necessary for the development of these equations. Next, we will introduce the concept of Wiener process together with its properties. The objective will be to build a Wiener process that will allow us to model random components for such differential equations. Next, we will present the definition and properties of the Itô stochastic integral which will lead us to the concept of solution of a stochastic differential equation, establishing the existence and uniqueness of such solution. Finally, we will see some examples and applications of numerical methods such as the Euler-Maruyama and the Milstein methods.

1. Introduction

In this work we will study the stochastic differential equation (SDE)

$$\begin{cases} dX = b(t, X)dt + B(t, X)dW \\ X(0) = X_0 \end{cases}, \quad 0 \leq t \leq T, \quad (1)$$

The stochastic differential equations are differential equations which contain a deterministic part and an additional random term defined in terms of a Wiener process. They are used to model and study various random phenomena. In having random terms, the classical calculus is not adequate, so it is necessary a introduction of an stochastic calculus which gives sense to such models and to the solutions of such differential equations.

Definition 1 An stochastic process in a probability space (Ω, \mathcal{U}, P) is an application $X: [0, +\infty) \times \Omega \rightarrow \mathbb{R}^n$ that $X(t, \cdot) \forall t \in [0, +\infty)$ is a random variable. The application $X(\cdot, \omega)$ given $\omega \in \Omega$ is a sample path.

2. Wiener processes

A real stochastic process W is called **Wiener process** if

- (i) $W(0) = 0$ almost sure.
- (ii) $W(t) - W(s) \text{ es } N(0, t-s), \forall t \geq s \geq 0.$
- (iii) $W(t) - W(s)$ y $W(v) - W(u)$ are independent for $0 \leq s < t < u < v.$

Theorem 2 Let $\{A_k\}_{k=1}^{\infty}$ be sequence of independent $N(0,1)$ random variables defined in the same space probabilistic

(Ω, \mathcal{U}, P) . Then there exists a Wiener process defined for all $\omega \in \Omega$ and $t \geq 0$. Moreover, the sample path $t \rightarrow X(t, \omega)$ are almost sure continuous.

3. Itô stochastic integrals

Lema 3 Let $[a, b]$ be an interval in $[0, \infty)$ and $\mathcal{P}^n := \{a = t_0^n < t_1^n < \dots < t_{m_n}^n = b\}$ partitions of $[a, b]$ with $|\mathcal{P}^n| \rightarrow 0, n \rightarrow \infty$. Then

$$\sum_{k=0}^{m_n-1} (W(t_{k+1}^n) - W(t_k^n))^2 \xrightarrow[n \rightarrow \infty]{L^2(\Omega)} b - a. \quad (2)$$

Remark 4 This means that $dW \approx (dt)^{\frac{1}{2}}$ in $L^2(\Omega)$ sense.

Lema 5 Let $G \in L^2(0, T)$. There exists a sequence of bounded step processes $G^n \in L^2(0, T)$ that

$$\lim_{n \rightarrow \infty} E \left[\int_0^T |G - G^n|^2 dt \right] \rightarrow 0. \quad (3)$$

Definition 6 Let $G \in L^2(0, T)$ and $\{G^n\}_{n=1}^{\infty} \subset L^2(0, T)$ be a sequence of bounded step processes fulfilling (3). Then, the Itô's integral of G is defined by

$$\int_0^T G dW := \lim_{n \rightarrow \infty} \sum_{k=0}^{m_n-1} G_k (W(t_{k+1}) - W(t_k)). \quad (4)$$

4. Stochastic Differential Equations

Definition 7 An stochastic process X in \mathbb{R}^n is a solution of the Itô stochastic differential equation (1) if $X(\cdot)$ fulfils almost surely for all $t \in [0, T]$

$$X(t) = X_0 + \int_0^t b(s, X(s)) ds + \int_0^t B(s, X(s)) dW. \quad (5)$$

Theorem 8 (Existence and Uniqueness) Let b and $B: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ be continuous functions fulfilling for all $0 \leq t \leq T, x, \hat{x} \in \mathbb{R}^n$

- a) $|b(t, x) - b(t, \hat{x})| \leq L|x - \hat{x}|$ and $|B(t, x) - B(t, \hat{x})| \leq L|x - \hat{x}|,$
- b) $|b(t, x)| \leq L(1 + |x|)$ and $|B(t, x)| \leq L(1 + |x|),$

for some constant L independent of T . Moreover, let X_0 be a random variable in \mathbb{R}^n fulfilling

- c) $E(|X_0|^2) < \infty$
- d) X_0 independent from $W^*(0)$

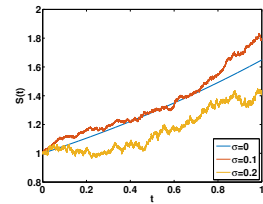
where $W(\cdot)$ is a one-dimensional Wiener process. Then, there exists a unique solution $X \in L^2_{\text{loc}}(0, T)$ for the SDE (1).

5. Numerical illustrations

Example 9 Geometric Brownian Motion

$$\begin{cases} \frac{dS}{S} = \mu dt + \sigma dW \\ S(0) = s_0 \end{cases} \quad (6)$$

with solution $S(t) = s_0 \cdot e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)}$.



Euler-Maruyama method

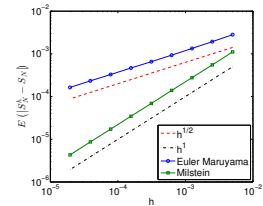
$$X_{n+1}^h = X_n^h + b(t_n, X_n^h) \cdot h_n + B(t_n, X_n^h) \cdot \Delta W_n$$

Milstein method

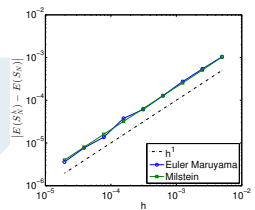
$$X_{n+1}^h = X_n^h + b(t_n, X_n^h)h_n + B(t_n, X_n^h)\Delta W_n + \frac{1}{2}(B' \cdot B)(t_n, X_n^h)((\Delta W_n)^2 - h_n)$$

with $\Delta W_n := W(t_{n+1}) - W(t_n) \sim N(0, h_n)$ independent random variables, and $h_n = t_{n+1} - t_n, 0 \leq n \leq N-1$.

Orden fuerte



Orden débil



References

- [1] L.C. Evans, *An Introduction to Stochastic Differential Equations*, AMS 2013.
- [2] P.E. Kloeden y E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer 1999.