Curso Avanzado de Probabilidad, 2014-I
Integración Estocástica

Gerónimo Uribe Bravo
Instituto de Matemáticas
Universidad Nacional Autónoma de México

Contenido

CAPÍTULO 1

Preliminares

1. Funciones de variación acotada y la integral de Lebesgue-Stieltjes

Definición. Decimos que una función $f : [0, \infty) \to \mathbb{R}$ es de **variación acotada en el intervalo** $[0, t]$ si

$$\sup_{\pi=\{0=t_0<\cdots<t_n=t\}} \sum_i |f(t_i) - f(t_{i-1})| < \infty,$$

donde el supremo es sobre todas las particiones de $[0, t]$. Denotaremos por $V_t(f)$ al supremo anterior.

Diremos que f tiene **variación localmente acotada** si $V_t(f) < \infty$ para toda $t \geq 0$ y que f tiene **variación acotada** si $\lim_{t\to\infty} V_t(f) < \infty$.

Si f es una función de variación localmente acotada, la célebre descomposición de Jordan nos afirma que la podemos escribir como diferencia de dos funciones no decrecientes. En efecto, es fácil verificar que

$$t \mapsto \frac{V_t(f) + f(t)}{2}, \frac{V_t(f) - f(t)}{2}$$

son no decrecientes y su suma es $f(t)$. Una forma intuitiva de entender la descomposición de Jordan se obtiene de definir la variación positiva o negativa de una función de variación acotada como

$$\sup_{\pi=\{0=t_0<\cdots<t_n=t\}} \sum_i |f(t_i) - f(t_{i-1})|^{\pm} \leq V_t(f) < \infty$$

y notar que ambas funciones son no decrecientes. Por otra parte, se tienen las descomposiciones

$$V_t^+(f) = \frac{V_t(f) + f(t) - f(0)}{2} \quad \text{y} \quad V_t^-(f) = \frac{V_t(f) - f(t) + f(0)}{2}.$$

Aún más, notemos que si f es también continua por la derecha entonces f se puede escribir como la diferencia de dos funciones no-decrecientes continuas por la derecha. En efecto, basta notar que si f es continua por la derecha en t entonces su variación también es continua en t. Esto implica también la existencia de límites por la izquierda. Puesto que cada una de estas funciones no decrecientes está asociada a una medida, vemos que existe una medida con signo μ tal que $f(t) = \mu([0, t])$. Si $g : [0, \infty) \to \mathbb{R}$ es localmente acotada y medible, podemos entonces definir a la **integral de Lebesgue-Stieltjes** de g respecto de f en $[0,t]$, denotada $\int_0^t g\,df$ mediante

$$\int_0^t g\,df = \int_0^t g\,d\mu.$$

El teorema siguiente nos afirma que las funciones de variación acotada aparecen naturalmente al tratar de considerar integrales elementales y extenderlas por "continuidad". Si nuestra función f es función continua por la derecha y π_n es una sucesión de particiones de $[0,t]$ que se refinan y cuyo paso tiende a cero, definamos

$$S^n(h) = \sum_{\pi_n} h(t_i)\left(f(t_{i+1}) - f(t_i)\right).$$

Si h es continua y f es de variación acotada en $[0,t]$ entonces $S^n(h) \to \int_0^t h\,df$.

Teorema 1.1. *La sucesión $(S^n(h), n \in \mathbb{N})$ converge a un límite para toda función continua h si y sólo si f es de variación acotada en $[0,t]$.*

Demostración. Ya hemos mostrado una de las implicaciones. Para la recíproca, supongamos que $(S^n(h), n \in \mathbb{N})$ converge a un límite para toda función continua h en $[0,t]$. Es bien sabido que el espacio de funciones continuas en $[0,t]$, al dotarlo de la norma uniforme, es un espacio de Banach. Además, S^n es una funcional lineal continua definido en él. Es fácil construir una función h de norma uniforme 1 tal que

$$S_n(h) = \sum_{\pi_n} |f(t_i) - f(t_{i-1})|$$

(por ejemplo al imponer que $h(t_i) = \mathrm{sgn}(f(t_i) - f(t_{i-1}))$ e interpolar linealmente). Por lo tanto,

$$V_t(f) \leq \sup_n \|S^n\|.$$

Por otra parte, para toda h continua se tiene que

$$\sup_n |S^n(h)| < \infty,$$

por lo que el principio de acotamiento uniforme (conocido también como teorema de Banach-Steinhaus) implica que

$$\sup_n \|S^n\| < \infty.$$

Así, vemos que f tiene variación acotada. □

Ejercicio 1.1. Pruebe que si f es no decreciente y continuamente diferenciable entonces es de variación actodada y

$$\int_0^t g\,df = \int_0^t g f'\,d\lambda.$$

Análogamente a como se definió la noción de variación de una función, podemos definir el de p-variación para $p > 0$. Este concepto será de utilidad al estudiar las propiedades trayectoriales del movimiento browniano (y otros procesos estocásticos).

DEFINICIÓN. Sean $f : [0, \infty) \to \mathbb{R}$ y $p > 0$. Se define la p-variación de f en $[0, t]$ como

$$\sup_{\pi=\{0=t_0<\cdots<t_n=t\}} \sum_i |f(t_i) - f(t_{i-1})|^p.$$

La denotaremos por $V_t^p(f)$.

PROPOSICIÓN 1.1. *Si f es una función continua de p-variación finita en $[0, t]$ entonces su p'-variación es cero si $p' > p$ e infinito si $p' < p$.*

DEMOSTRACIÓN. Sean π_n una sucesión de particiones cuyo paso tiende a cero y tales que

$$\sum_{\pi_n} |f(t_i) - f(t_{i-1})|^p \to V_t^p(f) < \infty,$$

Si $p < p'$, notemos que

$$\sum_{\pi_n} |f(t_i) - f(t_{i-1})|^{p'} \leq \max_{\pi_n} |f(t_i) - f(t_{i-1})|^{p'-p} \sum_{\pi_n} |f(t_i) - f(t_{i-1})|^p.$$

Como f es continua, el máximo del lado derecho tiende a cero, mientras que la suma del lado derecho tiene un límite finito. Por lo tanto, el lado izquierdo tiende a cero. Un razonamiento análogo prueba la afirmación restante. □

Concluimos la sección con algunas propiedades de la integral de Lebesgue-Stieltjes que contrastaremos con propiedades de la integral estocástica. Comenzamos con la regla de asociatividad, la fórmula de integración por partes y la regla de la cadena.

La regla de asociatividad nos dice que, como función del intervalo de integración, una integral de Lebesgue-Stieltjes es de variación acotada y permite reexpresar la integral respecto de ella. Formalmente:

PROPOSICIÓN 1.2. *Sea f una función de variación localmente acotada, g una función medible y acotada y definamos*

$$h(t) = \int_0^t g\, df.$$

Entonces h es de variación localmente acotada y si $\tilde{g}$ es medible y acotada entonces

$$\int_0^t \tilde{g}\, dh = \int_0^t \tilde{g} g\, df.$$

Con la notación $h = g \cdot f$, podemos escribir la fórmula de manera compacta:

$$\tilde{g} \cdot (g \cdot f) = (\tilde{g} g) \cdot f.$$

Pasemos ahora a la fórmula de integración por partes.

Proposición 1.3. *Sean f y g dos funciones de variación localmente acotada. Para toda $t > 0$ se tiene que*

$$f(t)\, g(t) = f(0)\, g(0) + \int_0^t g_-\, df + \int_0^t f\, dg.$$

Demostración. Sean μ_f y μ_g las medidas con signo asociadas a f y a g y $\mu = \mu_f \otimes \mu_g$. Entonces

$$\mu\big([0,t]^2\big) = f(t)\, g(t)\,.$$

Por otra parte, puesto que

$$[0,t]^2 = \{(0,0)\} \cup \{(r,s) : 0 \leq r \leq s \leq t\} \cup \{(r,s) : 0 \leq s < r \leq t\}\,,$$

podemos utilizar el teorema de Tonelli-Fubini para concluir que

$$\mu\big([0,t]^2\big) = f(0)\, g(0) + \int_0^t f\, dg + \int_0^t g_-\, df. \qquad \square$$

La fórmula de integración por partes admite la siguiente forma más simétrica:

$$f(t)\, g(t) = f(0)\, g(0) + \int_0^t f_-\, dg + \int_0^t g_-\, df + \sum_{s \leq t} \Delta f(s)\, \Delta g(s)\,.$$

Un caso, que resultará interesante contrastar con la integral estocástica, es

$$f(t)^2 = f(0)^2 + \int_0^t 2 f_-\, df,$$

válido cuando f es continua y de variación localmente acotada.

Ejercicio 1.2. Probar que si f es una función continua de variación acotada entonces g satisface

$$g(t) = x + \int_0^t g(s)\ f(ds)$$

si y sólo si

$$g(t) = x e^{f(t)}.$$

Sugerencia: utilice integración por partes para ver que $g e^{-f}$ es constante.

Aunque el siguiente resultado admite una versión para funciones discontinuas, nos concentraremos en el caso continuo.

Proposición 1.4 (Regla de la cadena). *Si f es una función continua de variación localmente acotada y F es de clase C_1 entonces $F \circ f$ es continua, de variación localmente acotada y*

$$F \circ f(t) = F \circ f(0) + \int_0^t F' \circ f \, df.$$

Demostración. La idea de la prueba es notar que si la fórmula es válida para F y para G y $\alpha, \beta \in \mathbb{R}$ entonces la fórmula también es válida para la función $\alpha \operatorname{Id} F + \beta G$. Puesto que la fórmula es válida para la función identidad, también lo será para polinomios y, al aproximar a F' (y por ende a F) uniformemente en $[0, t]$ por una sucesión de polinomios, podremos concluir la fórmula para toda F continuamente diferenciable.

Basta entonces probar que si

$$F \circ f(t) = F \circ f(0) + \int_0^t F' \circ f \, df.$$

entonces

$$f(t)\, F \circ f(t) = \int_0^t \operatorname{Id}(F' \circ f) + F \circ f \, df.$$

Sin embargo, de acuerdo a la fórmula de integración por partes, vemos que

$$\begin{aligned} f(t)\, F \circ f(t) &= \int_0^t F \circ f \, df + \int_0^t \operatorname{Id} \, dF \circ f \\ &= \int_0^t F \circ f \, d\lambda + \int_0^t \operatorname{Id} F' \circ f \, df \\ &= \int_0^t F \circ f + \operatorname{Id} F' \circ f \, df \\ &= \int_0^t (\operatorname{Id} F)' \circ f \, df. \end{aligned}$$

□

Ejercicio 1.3. Al extender la demostración anterior, pruebe la regla de la cadena

$$F \circ f(t) = F \circ f(0) + \int_{(0,t]} F' \circ f \, df + \sum_{\substack{s \leq t \\ \Delta f(s) \neq 0}} \left[F \circ f(s) - F \circ f(s-) - F' \circ f(s-)\, \Delta f(s)\right]$$

donde F es una función de clase C_1 y f es una función de variación localmente acotada.

Finalmente, analizaremos la llamada fórmula de cambio de variable, al introducir un concepto fundamental llamado inverso continuo por la derecha y estudiar la técnica de cambios de tiempo.

Sea $f : [0, \infty) \to [0, \infty]$ no decreciente y continua por la derecha (por lo que también admite límites por la izquierda). Definimos a f^{-1} mediante

$$f^{-1}(t) = \inf \{s \geq 0 : f(s) > t\}.$$

La idea de la construcción es que la gráfica de f^{-1} se obtiene al intercambiar los ejes en la gráfica de f: los intervalos de constancia de f se convierten en saltos de f^{-1} y los saltos de f se convierten en intervalos de constancia de f^{-1}.

Proposición 1.5. *La función f^{-1} es no decreciente, continua por la derecha, $f \circ f^{-1} \geq \mathrm{Id}$ y además*

$$f(t) = \inf \left\{s > 0 : f^{-1}(s) > t\right\}.$$

Si f es continua entonces $f \circ f^{-1} = \mathrm{Id}$.

Informalmente podemos escribir la fórmula para f en términos de f^{-1} como $\left(f^{-1}\right)^{-1} = f$.

Demostración. Si $t \leq \tilde{t}$ entonces

$$\{s \geq 0 : f(s) > t\} \supset \left\{s \geq 0 : f(s) > \tilde{t}\right\}.$$

Al tomar ínfimos obtenemos la monotonía de f^{-1}.

Si $t_n \downarrow t$ entonces

$$\{s \geq 0 : f(s) > t\} = \bigcup_n \{s \geq 0 : f(s) > t_n\}.$$

Al tomar ínfimos, concluimos que $f^{-1}(t) = \lim_n f^{-1}(t_n)$. En efecto, es claro que $f^{-1}(t) \leq \lim_n f^{-1}(t_n)$ y por definición de $f^{-1}(t)$, vemos que para toda $\varepsilon > 0$

$$t < f\left(f^{-1}(t) + \varepsilon\right),$$

por lo cual

$$t_n < f\left(f^{-1}(t) + \varepsilon\right)$$

para n suficientemente grande y por lo tanto

$$\lim_n f^{-1}(t_n) \leq f^{-1}(t) + \varepsilon.$$

Por definición de $f^{-1}(t)$, si $\varepsilon > 0$ se satisface

$$f\left(f^{-1}(t) + \varepsilon\right) > t$$

y por continuidad por la derecha de f obtenemos

$$f \circ f^{-1}(t) \geq t.$$

Si $f^{-1}(s) > t$ entonces $f(t) \leq s$. Por otra parte, si $\varepsilon > 0$, al ser f continua por la derecha, vemos que existe una vecindad derecha de t en la cual $f \leq f(t) + \varepsilon$. Por lo tanto $f^{-1}(f(t) + \varepsilon) > t$ y así vemos que

$$f(t) = \inf \left\{s \geq 0 : f^{-1}(s) > t\right\}.$$

Finalmente, si f es continua y no decreciente, entonces es sobre. Si (x_n) decrece estrictamente a x, existe (t_n) estrictamente decreciente (digamos con límite t) tal que $f(t_n) = x_n$. Por continuidad, se sigue que $f(t) = x$ y que a la derecha de t $f > x$. Vemos que entonces $f^{-1}(x) = t$ y que $f \circ f^{-1}(x) = f(t) = x$. □

Ahora podemos enunciar la fórmula de cambio de variable.

Proposición 1.6. *Si $g : [0, \infty) \to [0, \infty)$ es Borel-medible y f es no-decreciente entonces*

$$\int_{[0,\infty)} g\, df = \int_{[0,\infty)} g \circ f^{-1} \mathbf{1}_{f^{-1} < \infty}\, d\lambda.$$

Demostración. Se sigue del procedimiento de aproximación estándar de funciones medibles al notar que la proposición es válida para $g = \mathbf{1}_{[0,t]}$; en efecto, basta notar que $\left\{x \geq 0 : f^{-1}(x) \leq t\right\}$ es un intervalo de tamaño $f(t)$. □

Si f es una función de clase C_1 con derivada estríctamente positiva, la fórmula anterior se reduce a

$$\int_0^t g(t)\, f'(t)\, dt = \int_0^{f(t)} g\big(f^{-1}(x)\big)\, dx.$$

Esto explica el nombre de fórmula de cambio de variable.

2. Espacios de Hilbert y proyecciones ortogonales

Definición. Sea H un espacio vectorial sobre $\mathbb{R}$. Un **producto interior** en H es una función $\langle \cdot\,, \cdot\rangle : H \times H \to \mathbb{R}$ que satisface las siguientes propiedades:

Positividad: Para cualquier $f \in H$ se tiene que $\langle f, f\rangle \geq 0$.

Definitividad: Si $\langle f, f\rangle = 0$ entonces $f = 0$.

Simetría: Para todo $f, g \in H$ se tiene que $\langle f, g\rangle = \langle g, f\rangle$.

Linealidad: Para todo $\alpha, \beta \in \mathbb{R}$ y $f, g, h \in H$ se tiene que

$$\langle \alpha f + \beta g, h\rangle = \alpha\langle f, h\rangle + \beta\langle g, h\rangle.$$

Si H es un espacio con producto interior, se satisface la desigualdad de Cauchy-Schwarz

$$\langle f, g\rangle \leq \langle f, f\rangle^{1/2} \langle g, g\rangle^{1/2}.$$

Por esta razón, la función $|\cdot| : H \to [0, \infty)$ dada por

$$|f| = \langle f, f\rangle^{1/2}$$

es una norma sobre H.

Definición. Un **espacio de Hilbert real** es un espacio vectorial H sobre $\mathbb{R}$ dotado de un producto interior $\langle\cdot,\cdot\rangle$ y norma $|\cdot|$ tal que si d es la distancia asociada a $|\cdot|$ entonces (H,d) es un espacio métrico completo.

A continuación trabajaremos con un espacio de Hilbert $(H,\langle\cdot,\cdot\rangle)$. Lo que separa a los espacios de Hilbert de otros espacios vectoriales normados (y completos) es la siguiente identidad del paralelogramo: para cualesquiera $f,g\in H$ se tiene que

$$|f+g|^2+|g-f|^2=2\left(|f|^2+|g|^2\right).$$

La identidad del paralelogramo es la componente principal en la construcción de proyecciones.

Proposición 1.7. *Sea M un subespacio cerrado de H. Para todo $f\in H$ existe un único $g\in M$ tal que*

$$d(f,M)=d(f,g).$$

El vector g admite la siguiente caracterización:

$$\langle f-g,h\rangle=0$$

para todo $h\in M$.

Demostración. Sea (g_n) una sucesión de elementos de M tales que

$$d(f,g_n)\to d(f,M).$$

Por la identidad del paralelogramo, aplicada a $(f-g_n)/2$ y $(f-g_m)/2$, se sigue que

$$d(g_n/2,g_m/2)^2+d(f,(g_n+g_m)/2)^2\leq\frac{1}{2}\left[d(f,g_n)^2+d(f,g_m)^2\right].$$

Sin embargo, puesto que M es un subespacio, se sigue que $(g_n+g_m)/2\in M$ y por lo tanto

$$0\leq d(g_n/2,g_m/2)^2\leq\frac{1}{2}\left[d(f,g_n)^2+d(f,g_m)^2\right]-d(f,M)^2.$$

Al tomar límites se obtiene

$$0=\lim_{m,n\to\infty}d(g_n/2,g_m/2).$$

La sucesión (g_n) es por lo tanto una sucesión de Cauchy y por completitud se sigue que converge. Sea $g=\lim_{n\to\infty}g_n$. Entonces

$$d(f,g)=\lim_{n\to\infty}d(f,g_n)=d(f,M).$$

Notemos que si g y $\tilde g$ fueran dos elementos de M tales que $d(f,M)=d(f,g)=d(f,\tilde g)$, la sucesión $g,\tilde g,g,\tilde g,\ldots$, al ser de Cauchy por el párrafo anterior, nos muestra que $g=\tilde g$.

Para verificar la caracterización propuesta para g notemos que si $d(f, M) = d(f, g)$ entonces

$$|f - g|^2 \leq |f - g \pm th|^2 = |f - g|^2 \pm \langle f - g, th \rangle + |th|^2,$$

para toda $h \in M$ y toda $t \in \mathbb{R}$ por lo que

$$\pm t\langle f - g, h \rangle \leq t^2 |h|^2.$$

Al dividir sobre $|t|$, escoger el múltiplo de h covenientemente, y tomar el límite conforme $t \to 0$ vemos que

$$\langle f - g, h \rangle = 0.$$

Finalmente, si g es cualquier elemento de M que satisface $\langle f - g, h \rangle = 0$ para toda $h \in M$ entonces

$$|f - h|^2 = |f - g|^2 + 2\langle h, g \rangle + |g - h|^2 = |f - g|^2 + |g - h|^2 \geq |f - g|^2$$

por lo que $d(f, M) = d(f, g)$. $\square$

El anterior teorema nos permitirá representar a las funcionales lineales continuas en espacios de Hilbert. Esta es la herramienta principal en la construcción de la integral estocástica. Como un ejemplo más sencillo, lo utilizaremos para dar una prueba de la existencia de la esperanza condicional.

Definición. Una **funcional lineal continua** es una aplicación lineal $\phi : H \to \mathbb{R}$ tal que

$$|\phi| := \sup_{\substack{h \in H \\ |h| \leq 1}} |\phi(H)\,| < \infty.$$

Teorema 1.2 (Teorema de representación de Riesz). *Sea ϕ una funcional lineal continua en H. Entonces existe un único $f \in H$ tal que $\phi(g) = \langle f, g \rangle$.*

Demostración. Sea $M = \phi^{-1}(0)$. Si $M = H$ entonces $\phi = 0$ y por lo tanto podemos tomar $f = 0$. A continuación supondremos que $M \neq H$. Claramente M es un subespacio de H; se afirma que es cerrado. En efecto, si (g_n) es una sucesión de M que converge a g entonces

$$|\phi(g)| = |\phi(g - g_n)| \leq |\phi||g - g_n| \to 0.$$

Por lo tanto $g \in M$.

Sea $\tilde{f} \in H \setminus M$. Sea $p_M\left(\tilde{f}\right)$ la proyección ortogonal de $\tilde{f}$ sobre M y definamos a

$$f = \frac{1}{\left|\tilde{f} - p_M\left(\tilde{f}\right)\right|}\left[\tilde{f} - p_M\left(\tilde{f}\right)\right].$$

Entonces

$$\langle f, f \rangle = 1 \quad \text{y} \quad \langle f, g \rangle = 0$$

para toda $g \in M$. Por otra parte, si $h \in H$, notemos que

$$\phi(h - \lambda f) = 0 \text{ si } \lambda = \frac{\phi(h)}{\phi(f)}.$$

Por lo tanto:

$$\langle h - \lambda f, f \rangle = 0$$

o equivalentemente:

$$\langle \phi(f)\, f, h \rangle = \phi(h)\,.$$

Se concluye que $\phi(f)\, f$ es un vector que satisface la conclusión del teorema.

Para mostrar la unicidad, notemos que si f y $\tilde{f}$ son tales que $\phi(h) = \langle f, h \rangle = \langle \tilde{f}, h \rangle$ para toda $h \in H$ entonces $\langle f - \tilde{f}, h \rangle = 0$ para toda $h \in H$ lo cual muestra que $f = \tilde{f}$. □

Un ejemplo de aplicación del resultado anterior, que es de especial interés al inicio de la teoría de la probabilidad, es mostrar la existencia de la esperenza condicional. En efecto, si $(\Omega, \mathscr{F}, \mathbb{P})$ es un espacio de probabibilidad y $\mathscr{G} \subset \mathscr{F}$ es una σ-álgebra entonces el espacio $L_2(\Omega, \mathscr{G}, \mathbb{P})$ es un espacio de Hilbert con producto escalar $(X, Y) \mapsto \mathbb{E}(XY)$. Además, si $X \in L_2(\Omega, \mathscr{F}, \mathbb{P})$, la asignación $Y \mapsto \mathbb{E}(XY)$ de $L_2(\Omega, \mathscr{G}, \mathbb{P})$ en $\mathbb{R}$ es una funcional lineal continua. Por lo tanto, existe $Z \in L_2(\Omega, \mathscr{G}, \mathbb{P})$ tal que $\mathbb{E}(XY) = \mathbb{E}(ZY)$ para toda $Y \in L_2(\Omega, \mathscr{G}, \mathbb{P})$. En particular, podemos tomar $Y = \mathbf{1}_A$ para cualquier $A \in \mathscr{G}$. Así, Z es una versión de la esperanza condicional de X dada $\mathscr{G}$. Una vez que se construye la esperanza condicional para variables aleatorias cuadrado integrables se puede extender por truncamiento y el teorema de convergencia dominada a variables aleatorias integrables.

3. Preliminares de procesos estocásticos

Un **proceso estocástico** es una colección de variables aleatorias

$$X = (X_t, t \in I)\,.$$

I será interpretado como un parámetro temporal y será ya sea $\mathbb{N}$ (tiempo discreto) ó $[0, \infty)$ (tiempo continuo) o algún subintervalo de estos. Cuando no se mencione, se asumirá que $I = [0, \infty)$.

Supondremos que cada X_t toma valores en el mismo espacio medible, llamado espacio de estados, que para nosotros será ya sea $\mathbb{R}$ ó $\mathbb{R}^n$ equipados con su σ-álgebra de Borel. La **trayectoria** del proceso estocástico es la función aleatoria que en cada $\omega \in \Omega$ está dada por $t \mapsto X_t(\omega)$.

Ahora definiremos 3 maneras en que dos procesos estocásticos X y Y pueden ser considerados iguales. Diremos que X y Y tienen las **mismas distribuciones finito dimensionales** si para cada $n \geq 1$ y $t_1, \ldots, t_n \in I$ se tiene la igualdad en distribución de los vectores aleatorios siguientes:

$$(X_{t_1}, \ldots, X_{t_n}) \stackrel{d}{=} (Y_{t_1}, \ldots, Y_{t_n})\,.$$

Por ejemplo, dos movimientos brownianos siempre tienen las mismas distribuciones finito-dimensionales. La simetría del movimiento browniano implica que si B es un movimiento browniano entonces $-B$ también lo es. Así, esta noción no distingue mucho entre procesos cuyas trayectorias pueden ser muy distintas.

Diremos que X es una **modificación** de Y si $\mathbb{P}(X_t = Y_t) = 1$ para toda $t \in I$. Por ejemplo, si U es una variable aleatoria uniforme en $[0, 1]$, $X_t = 0$ para toda $t \geq 0$ y $Y_t = \mathbf{1}_{U=t}$, entonces $\mathbb{P}(X_t = Y_t) = \mathbb{P}(U \neq t) = 1$. Así, X es modificación de Y. Es decir, el concepto de modificación ya permite distinguir entre procesos con trayectorias distintas: B no es una modificación de $-B$. Sin embargo, el concepto de modificación no es tan bueno para distinguir entre trayectorias discontinuas.

Un concepto aún más fuerte es el de indistinguibilidad. Decimos que X y Y son indistinguibles si $X_t = Y_t$ para toda $t \geq 0$ casi seguramente. Es decir, existe $N \in \mathscr{F}$ tal que $\mathbb{P}(N) = 0$ para el cual $X_t(\omega) = Y_t(\omega)$ para toda $t \geq 0$ si $\omega \in N^c$. En el ejemplo de dos procesos que son modificación, vemos que 0 y $\mathbf{1}_{U=t}$ ya no son indistinguibles. Sin embargo, cuando las trayectorias de dos procesos estocásticos son regulares, entonces los conceptos de modificación e indistinguibilidad son equivalentes.

Proposición 1.8. *Sean X y Y dos procesos estocásticos cuyas trayectorias son casi-seguramente continuas por la derecha. Si X es modificación de Y entonces X y Y son indistinguibles.*

También bastaría tener trayectorias continuas por la izquierda casi seguramente.

Demostración. Sea $S \subset I$ denso y numerable y N_1 el complemento de $\{X_t = Y_t \forall t \in S\}$. Por σ-subaditividad, vemos que

$$\mathbb{P}(N_1) \leq \sum_{t \in S} \mathbb{P}(X_t \neq Y_t) = 0.$$

Sea $N_2 \in \mathscr{F}$ de probabilidad cero tal que las trayectorias de X y de Y son continuas por la derecha en N_2^c. Definamos $N = N_1 \cup N_2$. Entonces $N \in \mathscr{F}$, $\mathbb{P}(N) = 0$ y en el complemento de N, tenemos primero que $X_t = Y_t$ para toda $t \in S$ y, puesto que X y Y tienen trayectorias continuas por la derecha y S es denso, vemos que si $t \in I$:

$$X_t = \lim_{s \downarrow t, s \in S} X_s = \lim_{s \downarrow t, s \in S} Y_s = Y_t,$$

por lo que en N^c se tiene que $X_t = Y_t$ para todo $t \in I$. $\square$

Esta es una razón por la cual preferimos a procesos estocásticos con trayectorias regulares. Muchos procesos que ocurren en la práctica, como los procesos de Markov que satisfacen la propiedad de Feller, tienen una modificación con trayectorias regulares. Esto es consecuencia del teorema de regularización de martingalas que veremos más tarde.

Utilizaremos el parámetro temporal para introducir el flujo del tiempo en el espacio de probabilidad. Esto se hace a partir del concepto de filtración. Una **filtración** es una colección $(\mathscr{F}_t, t \in I)$ de subσ-álgebras de $\mathscr{F}$ tales que si $t_1 \leq t_2 \in I$ entonces $\mathscr{F}_{t_1} \subset \mathscr{F}_{t_2}$. A la σ-álgebra $\mathscr{F}_t$ se le interpreta como la información disponible al tiempo t. Decimos que la filtración es **continua por la derecha** si

$$\mathscr{F}_t = \bigcap_{\varepsilon > 0} \mathscr{F}_{t+\varepsilon}.$$

Decimos que la filtración es **completa** si $\mathscr{F}_0$ contiene a todos los conjuntos $\mathbb{P}$ nulos. Se dice que una filtración satisface las hipótesis habituales si es continua por la derecha y completa.

Para impedir que nuestros procesos estocásticos requieran información futura en su evolución se utiliza el concepto de proceso adaptado. Un proceso estocástico X es **adaptado** a la filtración $(\mathscr{F}_t)$ si X_t es $\mathscr{F}_t$-medible para toda $t \geq 0$. Hay una filtración muy sencilla que hace adaptado a un proceso estocástico X: la **filtración canónica** $\left(\mathscr{F}_t^X, t \in I\right)$ dada por $\mathscr{F}_t^X = \sigma(X_s : s \leq t)$.

Hay un concepto más fuerte de medibilidad posible para procesos estocásticos que toma en cuenta el parámetro temporal. Decimos que un proceso estocástico X es **medible** si la función $(t, \omega) \mapsto X_t(\omega)$ de $I \times \Omega$ en el espacio de estados es medible.

PROPOSICIÓN 1.9. *Sea X un proceso estocástico con trayectorias continuas por la derecha. Entonces X es un proceso medible.*

DEMOSTRACIÓN. Por definición de σ-álgebra producto, la función

$$(t, \omega) \mapsto X_{\lceil 2^n t \rceil / 2^n}(\omega)$$

es $(\mathscr{B}_T \otimes \mathscr{F}, \mathscr{B}_{\mathbb{R}})$-medible. Puesto que su límite puntual conforme $n \to \infty$ es $(t, \omega) \mapsto X_t(\omega)$, vemos que X es un proceso medible. $\square$

Decimos que un proceso estocástico es **progresivamente medible** si la aplicación $(t, \omega) \mapsto X_t(\omega)$ de $[0, t_0] \times \Omega$ en el espacio de estados es $\mathscr{B}_{[0,t_0]} \otimes \mathscr{F}_{t_0}$-medible para toda $t_0 \geq 0$. Copiando la última prueba se verifica que si un proceso es adaptado y con trayectorias continuas por la derecha entonces es progresivamente medible.

Nuestro último preliminar de la teoría general de procesos será el concepto de tiempo de paro. Comenzamos con la definición de tiempo aleatorio:

DEFINICIÓN. Un **tiempo aleatorio** es una variable aleatoria extendida $T : \Omega \to [0, \infty]$.

Dado un tiempo aleatorio T finito y un proceso estocástico X, podemos formar a la función $X_T : \Omega \to \mathbb{R}$ dada por

$$X_T(\omega) = X_{T(\omega)}(\omega).$$

Si X es un proceso medible entonces X_T es una variable aleatoria. Mediante el concepto de tiempo de paro se puede extender este resultado para incorporar el flujo del tiempo.

Definición. Un **tiempo de paro** respecto de una filtración $(\mathscr{F}_t, t \in I)$ es una función $T : \Omega \to [0, \infty]$ tal que para toda $t \in I$: $\{T \leq t\} \in \mathscr{F}_t$. Un **tiempo opcional** es una función $T : \Omega \to [0, \infty]$ tal que para toda $t \in I$: $\{T < t\} \in \mathscr{F}_t$.

Cualquier tiempo de paro es un tiempo aleatorio. Claramente, cualquier tiempo de paro es un tiempo opcional. A veces, es más sencillo probar que un determinado tiempo aleatorio es opcional. En una filtración continua por la derecha, un tiempo opcional es un tiempo de paro. En general este no es el caso.

A continuación damos algunos ejemplos importantes de tiempos de paro.

Sea X un proceso estocástico y sea A un subconjunto del espacio de estados. Definamos

$$\mathscr{A} = \{t \geq 0 : X_t \in A\} \quad \text{y} \quad H_A = \begin{cases} \infty & \mathscr{A} = \emptyset \\ \inf \mathscr{A} & \mathscr{A} \neq \emptyset \end{cases}.$$

A H_A se le conoce como tiempo de arribo de X al conjunto A. Introduciremos ahora la convención $\inf \emptyset = \infty$. Co esta convención, escribimos de manera más compacta la definición de H_A como sigue:

$$H_A = \inf \{t \geq 0 : X_t \in A\}.$$

Proposición 1.10. *Si X tiene trayectorias continuas y A es cerrado entonces H_A es un tiempo de paro respecto de $\left(\mathscr{F}_t^X, t \geq 0\right)$. Si X tiene trayectorias continuas por la derecha, A es un conjunto abierto, la filtración $(\mathscr{F}_t)$ es continua por la derecha y X es adaptado entonces H_A es un tiempo de paro.*

Demostración. En toda generalidad podemos escribir

$$\{H_A \leq t\} = \{H_A = t\} \cup \{\exists s \leq t : X_s \in A\},$$

donde además

$$\{H_A = t\} = \{X_s \notin A \forall s < t, \forall \varepsilon > 0 \exists t_\varepsilon \in [t, t + \varepsilon], X_{t_\varepsilon} \in A\}.$$

Por lo tanto, el hecho de que H_A sea tiempo de paro se debe a que, con hipótesis adecuadas, se pueden discretizar las uniones e intersecciones.

En el primer caso vemos que $\{H_A = t\} \subset \{X_t \in A\}$ por lo que:

$$\begin{aligned} &\{H_A \leq t\} \\ &= \{\exists s \leq t : X_s \in A\} \\ &= \bigcap_{n \geq 1} \{\exists s \leq t : d(X_s, A) < 1/n\} \\ &= \bigcap_{n \geq 1} \bigcup_{s \in [0,t] \cap \mathbb{Q}} \{d(X_s, A) < 1/n\} \in \mathscr{F}_t^X \end{aligned}$$

En el segundo caso, basta probar que H_A es opcional puesto que la filtración es continua por la derecha. Por otra parte, puesto que las trayectorias de X son continuas por la derecha y A es abierto, vemos que si $X_s \in A$ para alguna $s < t$ entonces necesariamente $X_{s'} \in A$ para alguna $s' \in [0,t) \cap \mathbb{Q}$, por lo que

$$\{H_A < t\} = \bigcap_{s \in [0,t) \cap \mathbb{Q}} \{X_s \in A\} .$$

Puesto que X es adaptado, el lado derecho pertenece a $\mathscr{F}_t$. □

Un resultado aún mucho más profundo, aunque muchísimo más difícil de demostrar es el siguiente:

Teorema 1.3. *Si A es Borel medible y X es un proceso progresivo respecto de una filtración càd y completa $(\mathscr{F}_t, t \geq 0)$ entonces H_A es un tiempo de paro.*

Proposición 1.11. *Si S y T son tiempos de paro entonces $S \wedge T$, $S \vee T$ y $S + T$ son tiempos de paros. Si $T_n, n \geq 1$ es una sucesión de tiempos de paro entonces $\sup_n T_n$ es un tiempo de paro y si la sucesión es creciente entonces $\lim_n T_n$ es un tiempo de paro.*

Demostración. Ejercicio. □

Definición. Sea T un tiempo de paro respecto a la filtración $(\mathscr{F}_t, t \geq 0)$. Definimos a la σ-álgebra de eventos anteriores a T como la σ-álgebra $\mathscr{F}_T$ dada por

$$\mathscr{F}_t = \{A \in \mathscr{F} : A \cap \{T \leq t\} \in \mathscr{F}_t\} .$$

Proposición 1.12. *La σ-álgebra de eventos anteriores a T es una σ-álgebra. Si $A \in \mathscr{F}_S$ entonces $A \cap \{S \leq T\} \in \mathscr{F}_T$ y en consecuencia, si $S \leq T$ entonces $\mathscr{F}_S \subset \mathscr{F}_T$. Además, $\mathscr{F}_{T \wedge S} = \mathscr{F}_T \cap \mathscr{F}_S$ y dicha σ-álgebra contiene a cada uno de los eventos*

$$\{S < T\}, \{T < S\}, \{S \leq T\}, \{T \leq S\}, \{T = S\}$$

Proposición 1.13. *Sea $X = (X_t, t \geq 0)$ un proceso estocástico progresivamente medible respecto de una filtración $(\mathscr{F}_t, t \geq 0)$. Si T es un tiempo de paro finito entonces la variable aleatoria X_T es $\mathscr{F}_T$-medible. Además, el proceso detenido X^T donde $X^T_t = X_{t \wedge T}$ es progresivamente medible.*

Ejercicio 1.4. Problema 1.2.19 p. 9 de Karatzas.

Ejercicio 1.5 ((Básicamente) Problema 1.2.21 p.10 de [**KS91**]). Sea T un tiempo opcional y defina la σ-álgebra de eventos que suceden hasta inmediatamente después de T como

$$\mathscr{F}_{T+} = \{A \in \mathscr{F} : A \cap \{T \leq t\} \in \mathscr{F}_{t+} \text{ para toda } t \geq 0\} .$$

Pruebe que $\mathscr{F}_{T+}$ es una σ-álgebra, que T es $\mathscr{F}_{T+}$-medible y que

$$\mathscr{F}_{T+} = \{A \in \mathscr{F} : A \cap \{T < t\} \in \mathscr{F}_t\} .$$

Pruebe que si T es tiempo de paro entonces $\mathscr{F}_T \subset \mathscr{F}_{T+}$ y que si la filtración es continua por la derecha entonces $\mathscr{F}_{T+} = \mathscr{F}_T$.

4. Martingalas a tiempo continuo

Comenzaremos recordando algunos resultados de la teoría de martingalas que daremos por hecho. Comencemos por la definición. Recordemos que $T \subset [0, \infty)$ representa el conjunto de tiempos que estamos considerando y que en general tomaremos como $[0, \infty)$, $\mathbb{N}$ ó algún intervalo de la forma $[0, t]$ ó $\{0, \ldots, n\}$.

Definición. Sean $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad y $(\mathscr{F}_t)_{t \in T}$ una filtración en dicho espacio. Una colección de variables aleatorias reales $(X_t)_{t \in T}$ es una martingala respecto a la filtración considerada si

(1) X_t es $\mathscr{F}_t$-medible,
(2) $X_t \in L_1$ y
(3) si $s \leq t$ entonces $\mathbb{E}(X_t \mid \mathscr{F}_s) = X_s$.

Una submartingala es un proceso estocástico que satisface 1 y 2 de la definición de martingala y en vez de 3 se satisface la desigualdad $\mathbb{E}(X_t \mid \mathscr{F}_s) \geq X_s$. En una supermartingala, la desigualdad será invertida.

Un ejemplo fundamental es la transformación de martingalas que permite obtener martingalas a partir de otras. Sea $M = (M_n, n \in \mathbb{N})$ una martingala respecto de $(\mathscr{F}_n, n \in \mathbb{N})$ y $C = (C_i, i \in \mathbb{Z}_+)$ un **proceso predecible** y acotado; esto es, C_i es $\mathscr{F}_{i-1}$-medible. Se define a la transformada de martingala $C \cdot M$ como el proceso estocástico dado por

$$(C \cdot M)_n = \sum_{i=1}^{n} C_i \left(M_i - M_{i-1}\right).$$

Ejercicio 1.6. Sea $(\mathscr{F}_n, n \in \mathbb{N})$ una filtración y $M = (M_n, \in \mathbb{N})$ una $(\mathscr{F}_n)$-martingala. Sea $C = (C_i, i \geq 1)$ un proceso acotado, digamos por K. Asuma que C_n es $\mathscr{F}_{n-1}$-medible.

(1) Defina al proceso estocástico $C \cdot M$ mediante

$$(C \cdot M)_0 = 0 \quad \text{y} \quad (C \cdot M)_n = \sum_{i=1}^{n} C_i \left(M_i - M_{i-1}\right) \text{ para } n \geq 1.$$

Pruebe cuidadosamente que $C \cdot M$ es una $(\mathscr{F}_n)$-martingala.

(2) Encuentre una cota para el segundo momento de $C \cdot M$ en términos del segundo momento de M y de K. Escriba la desigualdad L_2 de Doob para $C \cdot M$ y reexprésela con la cota que obtuvo.

Los resultados principales que utilizaremos sobre las (sub o super)martingalas son: el teorema de muestreo opcional de Doob, el lema de cruces de Doob, los teoremas de convergencia casi segura y en L_p, los teoremas de regularización y las desigualdades maximales. A continuación se enuncian.

Teorema 1.4. *Sea $(M_t, t \geq 0)$ una martingala cerrada (con trayectorias continuas por la derecha en el caso de tiempo continuo) respecto de $(\mathscr{F}_t, t \geq 0)$ con último elemento M_∞. Si $S \leq T$ son tiempos de paro entonces*

$$\mathbb{E}(M_T \mid \mathscr{F}_{S+}) = M_S.$$

Ejercicio 1.7. En el contexto del teorema anterior, pruebe que $\mathbb{E}(M_T \mid \mathscr{F}_S) = M_S$. Pruebe que obviamente $M_{t\wedge T}$ es una $\mathscr{F}_{t\wedge T}$-martingala pero que de hecho M^T es martingala respecto a la filtración $(\mathscr{F}_t)$.

Demostración. Se supondrá el resultado cuando $I = \mathbb{N}$ y lo demostraremos cuando $I = [0, \infty)$ y la martingala M tiene trayectorias càd. Definimos a

$$S^n = \lceil 2^n S \rceil / 2^n \quad \text{y} \quad T^n = \lceil 2^n T \rceil / 2^n.$$

Entonces S^n y T^n son dos tiempos aleatorios que toman valores en los racionales diádicos de orden n. Son además tiempos de paro respecto de la filtración a tiempo discreto $\left(\mathscr{F}_{k/2^n}, k \in \mathbb{N}\right)$ y se tiene que $S^n \leq T^n$, S^n a S y T^n decrece a T. Puesto que $\left(M_{k/2^n}, k \in \mathbb{N}\right)$ es una martingala respecto a la filtración a tiempo discreto que hemos definido vemos que

$$\mathbb{E}(M_{T^n} \mid \mathscr{F}_{S^n}) = M_{S^n}.$$

Puesto que M es uniformemente integrable, al ser cerrada, la continuidad por la derecha de las trayectorias de M nos dice que $M_{T^n} \to M_T$ casi seguramente y en L_1. Por otra parte, la colección decreciente de σ-álgebras $(\mathscr{F}_{S^n}, n \in \mathbb{N})$ tiene intersección igual a $\mathscr{F}_{S+}$. Por el teorema de convergencia hacia abajo de Lévy se sigue entonces que

$$\lim_{n\to\infty} \mathbb{E}(M_{T^n} \mid \mathscr{F}_{S^n}) = \mathbb{E}\big(M^T \bigm| \mathscr{F}_{S+}\big)$$

casi seguramente y en L_1. Puesto que $M_{S^n} \to M_S$ casi seguramente y en L_1 se sigue que

$$\mathbb{E}(M_T \mid \mathscr{F}_{S+}) = M_S. \qquad \square$$

Si $a < b$ son dos racionales, veamos cuantas veces cruzan hacia arriba los puntos $x_0, x_1, \ldots$ a $[a, b]$, cantidad que denotamos por $U_{[a,b]}(x_0, x_1, \ldots)$ y cuyo cálculo procedemos a explicar: sean

$$A_1 = \{k \in \mathbb{N} : x_k \leq a\},$$

$$T_1(x_0, x_1, \ldots) = \begin{cases} \min A_1 & \text{si } A_1 \neq \emptyset \\ \infty & \text{si } A_1 = \emptyset \end{cases}$$

y de manera recursiva, para $j \geq 1$

$$A_{2j} = \{k \in \mathbb{N} : T_{2j-1} \leq k, x_k \geq b\},$$

$$T_{2j}(x_0, x_1, \ldots) = \begin{cases} \min A_{2j} & \text{si } A_{2j} \neq \emptyset \\ \infty & \text{si } A_{2j} = \emptyset \end{cases}$$

$$A_{2j+1} = \{k \in \mathbb{N} : T_{2j} \leq k, x_k \leq a\}$$

$$T_{2j+1}(x_0, x_1, \ldots) = \begin{cases} \min A_{2j+1} & \text{si } A_{2j+1} \neq \emptyset \\ \infty & \text{si } A_{2j+1} = \emptyset \end{cases}.$$

A partir de las anteriores cantidades, definimos

$$\begin{aligned} U_{[a,b]}(x_0, \ldots, x_n) &= \sup\{k \in \mathbb{N} : k \geq 1, A_{2k} \neq \emptyset\} \\ &= \sup\{k \in \mathbb{N} : k \geq 1, T_{2k} < \infty\}, \end{aligned}$$

que es la cantidad de cruces hacia arriba de la sucesión en $[a, b]$ pues si que $T_{2k} < \infty$ entonces la sucesión ha cruzado $[a, b]$ hacia arriba de menos k veces, por definición de T_{2k}.

LEMA 1. *El límite inferior de la sucesión $(x_n)_{n\in\mathbb{N}}$ coincide con el límite superior de la misma si y sólo si para cualquier pareja de racionales $a < b$, la cantidad $U_{[a,b]}(x_0, x_1, \ldots)$ es finita. Si $(X_n, n \in \mathbb{N})$ es una martingala entonces*

$$\mathbb{E}\Big(U^n_{[a,b]}\Big) \leq \frac{1}{b-a}\mathbb{E}\Big((a - X_n)^+\Big).$$

A la desigualdad entre esperanzas se le conoce como la desigualdad de cruces de Doob.

TEOREMA 1.5 (Teorema fundamental de convergencia de martingalas). *Si (X_n) es una martingala acotada en L_1 entonces $(X_n)_{n\in\mathbb{N}}$ converge casi seguramente a una variable aleatoria X que pertenece a L_1.*

Ahora daremos una extensión adicional del teorema de convergencia de martingalas y probaremos el teorema de regularización de martingalas. Este último es útil a la construcción de una gran familia de procesos estocásticos entre los cuales se encuentran los procesos de Feller y en particular los procesos de Lévy. Nos centraremos en procesos a tiempo continuo.

Extenderemos ahora la noción de cantidad de cruces de una función $f : [0, \infty) \to \mathbb{R}$: recordemos que si $F \subset [0, \infty)$ es finito, ya tenemos definida la noción de la cantidad de cruces hacia arriba de $(f(t))_{t\in F}$ en el intervalo $[a, b]$, llamémosle $U_F(f, a, b)$. Si $T \subset \mathbb{R}$ es arbitrario, podemos definir

$$U_T(f, a, b) = \sup_{F \subset T, F \text{ finito}} U_F(f, a, b).$$

Es claro que si T es numerable y X es un proceso estocástico entonces $U_T(X, a, b)$ es una variable aleatoria. Por otra parte, si $T = [u, v] \cap \mathbb{Q}$, entonces para todo

$t \in [u, v]$ existen los límites

$$f(t+) = \lim_{s \downarrow t, s \in T} \quad t \in [u, v)$$

y

$$f(t-) = \lim_{s \uparrow t, s \in T} \quad t \in (u, v]$$

(y son reales) si y sólo si $U_T(f, a, b) < \infty$ para cualquier pareja de racionales $a < b$.

Teorema 1.6 (Desigualdad de cruces de Doob). *Si $(X_t)_{t \geq 0}$ es una $(\mathscr{F}_t)$-supermartingala càd entonces*

$$\mathbb{E}(U_T(X)) \leq \sup_{t \in T} \mathbb{E}\big((a - X_t)^-\big).$$

Nuestro objetivo ahora será demostrar la siguiente proposición.

Teorema 1.7. *Sea $(X_t, t \geq 0)$ una martingala respecto a una filtración $(\mathscr{F}_t, t \geq 0)$ continua por la derecha y débilmente completa. Entonces existe una modificación Y de X que también es una martingala respecto de $(\mathscr{F}_t, t \geq 0)$ y tal que Y tiene trayectorias càdlàg casi seguramente.*

Demostración. Para cualquier $n \in \mathbb{N}$ y $a < b$, la desigualdad de cruces de Doob nos dice que

$$\mathbb{E}\big(U_{[0,n] \cap \mathbb{Q}}(X, a, b)\big) \leq |a| + \mathbb{E}(|X_n|) < \infty,$$

por lo cual

$$\mathbb{P}\big(U_{[0,n] \cap \mathbb{Q}}(X, a, b) < \infty\big) = 1.$$

Por σ-subaditividad, vemos que

$$\mathbb{P}\big(U_{[0,n] \cap \mathbb{Q}}(X, a, b) < \infty \text{ si } a, b \in \mathbb{Q},\ a < b \text{ y } n \in \mathbb{N}\big) = 1.$$

En dicho conjunto, que denotaremos por N^c, X admite límites por la izquierda y por la derecha en t para todo $t \geq 0$ y por lo tanto podemos definir a

$$\tilde{X}_t(\omega) = \begin{cases} X_{t+}(\omega) & \text{si } \omega \in N^c \\ 0 & \text{si } \omega \in N \end{cases}.$$

Como X_{t+} es $\mathscr{F}_{t+}$-medible y $\mathscr{F}_{t+} = \mathscr{F}_t$ entonces X_{t+} es $\mathscr{F}_t$-medible y puesto que N pertenece a $\mathscr{F}_\infty$ y tiene probabilidad cero, entonces $N \in \mathscr{F}_t$ y por lo tanto $\tilde{X}_t$ es $\mathscr{F}_t$-medible. Por otra parte, en N^c, $\tilde{X}$ admite límites por la izquierda. Además, $\tilde{X}$ es continuo por la derecha en N^c por el argumento siguiente: si $\varepsilon > 0$ entonces existe $\delta > 0$ tal que si $r \in [t, t+\delta]$ y $\omega \in N^c$ entonces $\left|\tilde{X}_t(\omega) - X_r(\omega)\right| < \varepsilon$. Vemos que entonces, si $s \in [t, t+\delta)$ entonces $\left|\tilde{X}_t(\omega) - \tilde{X}_s(\omega)\right| \leq \varepsilon$.

Si $t_1 < t_2$ y s_n es una sucesión de racionales que decrecen a t_1, sabemos que

$$\mathbb{E}(\, X_{t_2} \mid \mathscr{F}_{s_n}) = X_{s_n}$$

y por el teorema de convergencia de Lévy hacia abajo, vemos que casi seguramente y en L_1:

$$\tilde{X}_{t_1} = \lim_{n\to\infty} X_{s_n} = \lim_{n\to\infty} \mathbb{E}(X_{t_2} \mid \mathscr{F}_{s_n}) = \mathbb{E}(X_{t_2} \mid \mathscr{F}_{t_1+}) = \mathbb{E}(X_{t_2} \mid \mathscr{F}_{t_1}) = X_{t_1},$$

por lo que $\tilde{X}$ es una modificación de X.

Consideremos ahora $t_1 < t_2$ y s_n una sucesión de racionales que decrezcan a t_2. Puesto que X_{s_n} converge casi seguramente y en L_1 a $\tilde{X}_{t_2}$, como vimos en el párrafo anterior, el teorema de convergencia dominada para la esperanza condicional nos dice que

$$\tilde{X}_{t_1} = \mathbb{E}(X_{s_n} \mid \mathscr{F}_{t_1}) \to \mathbb{E}\left(\tilde{X}_{t_2} \,\middle|\, \mathscr{F}_{t_1}\right),$$

por lo que $\tilde{X}$ es una $\mathscr{F}_t$-martingala. □

Ahora veremos un criterio sencillo para verificar que una martingala converge no sólo en L_1 sino también en L_p para algún $p > 1$. Para esto, estudiaremos al máximo valor que toma una martingala. Al contar con una cota adecuada para su esperanza máximo, se puede entonces simplemente aplicar convergencia dominada para verificar la convergencia en L_p de la martingala en cuestión. Sea $M = (M_t, t \in T)$ una $(\mathscr{F}_t)$-submartingala càd. Definamos a

$$\overline{M}_t^+ = \sup_{s\leq t} M_s^+.$$

Proposición 1.14 (Desigualdad maximal de Doob). *Para toda $\lambda > 0$,*

$$\lambda \mathbb{P}\left(\overline{M}_t^+ > \lambda\right) \leq \mathbb{E}\left(M_t^+\right).$$

Proposición 1.15 (Desigualdad L_p de Doob). *Para cualquier $p \in (1, \infty)$:*

$$\|\overline{M}_t^+\|_p \leq \frac{p}{p-1} \|M_t^+\|_p.$$

Teorema 1.8. *Si M_n es una martingala con $\sup_n \mathbb{E}(|M_n|^p) < \infty$ para alguna $p > 1$, X_n converge casi seguramente y en L_p a una variable M_∞ y se tiene que*

$$M_n = \mathbb{E}(M_\infty \mid \mathscr{F}_n).$$

CAPÍTULO 2

Integración estocástica respecto del movimiento browniano

En este capítulo se introducirá el ejemplo fundamental y más sencillo de integración estocástica: la integral estocástica respecto del movimiento browniano. Además, se introducirán y recordarán algunos conceptos de la teoría general de procesos estocásticos y de martingalas a tiempo continuo.

El proceso estocástico más importante en estas notas es el movimiento browniano. Recordemos su

DEFINICIÓN. Un **movimiento browniano** es una colección de variables aleatorias $B = (B_t, t \geq 0)$ tales que:

(1) B comienza en cero
(2) B tiene trayectorias continuas
(3) B tiene incrementos independientes y estacionarios
(4) La distribución de B_t es normal centrada de varianza t.

El teorema límite central nos permite entender a la distribución normal como una forma de aproximar a la suma de muchas contribuciones independientes (e idénticas desde el punto de vista probabilístico). Así, cada incremento del movimiento browniano se podría interpretar como una perturbación aleatoria obtenida de sumar muchas contribuciones pequeñas. Así, una forma de agregarle una fuente de error a una ecuación diferencial del tipo

$$dc_t = \varphi(c_t)\ dt$$

es considerar a

$$dc_t = \varphi(c_t)\ dt + dB_t \quad \text{ó equivalentemente} \quad c_t = x + \int_0^t \varphi(c_s)\ ds + B_t.$$

Un ejemplo muy concreto es el del proceso de Ornstein-Uhlenbeck, en el que $b(X_s) = \lambda X_s$. La interpretación en este caso es, si B denota nuestra ganancia o pérdida al jugar (continuamente) con un capital inicial x, X_t será nuestra ganancia si se paga interés a tasa λ por un préstamo ó un impuesto de la misma tasa.

Las ecuaciones diferenciales estocásticas surgen de la idea de hacer que la magnitud de la perturbación aleatoria dependan de la posición del sistema para llegar a ecuaciones del tipo

$$dX_t = b(X_t)\ dt + \sigma(X_t)\ dB_t \quad \text{ó} \quad X_t = \int_0^t b(X_s)\ ds + \int_0^t \sigma(X_s)\ dB_s.$$

Sin embargo, surge inmediatamente el problema de interpretar a la integral respecto del movimiento browniano. Puesto que, como veremos, las trayectorias de B tienen variación infinita en cualquier intervalo, entonces no se puede interpretar como una integral de Lebesgue-Stieltjes. Como veremos, aunque no sea posible definir a la integral estocástica como límite de sumas de Riemann trayectoria por trayectoria, sí se le puede definir como límite en probabilidad para una clase adecuada de integrandos (que no ven el futuro). La contribución de Itô fue darse cuenta de esto y entonces profundizar en el estudio de las similitudes y diferencias entre la integral usual y la integral estocástica. Al reinterpretar el artículo de Itô, algunos otros probabilistas como Doob se dieron cuenta de la similitud existente entre la integral estocástica y la transformada de martingala a tiempo discreto, lo cual marcó el desarrollo posterior del cálculo estocástico como una teoría apoyada fundamentalmente en las martingalas.

La integral estocástica respecto del movimiento browniano fue introducida en [**Itô87**]. En este artículo, Itô introduce la integral estocástica respecto del movimiento browniano con la idea de utilizarla para darle sentido a ecuaciones diferenciales estocásticas (para las cuales da un teorema de existencia y unicidad) y utilizar a estas para dar construcciones de procesos de Markov con trayectorias continuas. Seguiremos de cerca este artículo, comenzando con algunos preliminares del movimiento browniano.

1. El movimiento browniano y su variación cuadrática

Consideremos una caminata aleatoria simple y simétrica $S = (S_n, n \in \mathbb{N})$. El teorema límite central afirma que $S_n/\sqrt{n}$ converge débilmente a una variable normal estándar. Una manera de interpretar al movimiento browniano es como una extensión multidimensional (inclusive infinito-dimensional o funcional) del teorema límite central. En efecto, si S se extiende por interpolación lineal en cada intervalo $[n, n+1]$ y consideramos al proceso estocástico S^n dado por $S^n_t = S_{nt}/\sqrt{n}$, vemos que S^n_t converge débilmente a una normal de media 0 y varianza t. Por otra parte, como S tiene incrementos independientes y estacionarios (cuando nos restringimos a instantes de tiempo naturales) entonces si $0 = t_0 < t_1 < \cdots < t_m$ entonces para n suficientemente grande los incrementos $S^n_{t_i} - S^n_{t_{i-1}}$, con $1 \leq i \leq m$ son independientes. Por lo tanto, vemos que dichos incrementos convergen débilmente a un vector aleatorio con entradas gaussianas independientes de varianzas respectivas $t_i - t_{i-1}$ para $1 \leq i \leq m$. El movimiento browniano es justamente un proceso estocástico que recoge este comportamiento límite de las caminatas aleatorias.

Definición. Un **movimiento browniano en ley** es un proceso estocástico $B = (B_t, t \geq 0)$ tal que:

(1) $B_0 = 0$
(2) B tiene incrementos independientes: si $0 = t_0 < t_1 < \cdots < t_m$ entonces $B_{t_i} - B_{t_{i-1}}$, $1 \leq i \leq m$ son independientes
(3) B tiene incrementos estacionarios: $B_{t+s} - B_t$ tiene la misma distribución que B_s y
(4) la distribución de B_t es normal de media 0 y varianza t.

Un **movimiento browniano** es un movimiento browniano en ley que tiene trayectorias continuas.

No es trivial que el movimiento browniano exista. P. Lévy dió una demostración muy elegante de este hecho basada en propiedades de la distribución Gaussiana. Sin embargo, hay herramientas de aplicación más general y que permitían probar la existencia del movimiento browniano. En efecto, al aplicar el teorema de extensión de Kolmogorov, es fácil verificar la existencia de un espacio de probabilidad en el que está definido un movimiento browniano en ley. Al utilizar el criterio de continuidad de Kolmogorov se puede entonces encontrar una modificación del browniano en ley que sea un movimiento browniano. Ahora abordaremos otra prueba de existencia del browniano basado en el teorema de regularización de martingalas. No es una prueba ni tan corta ni tan elegante como las dos esbozadas anteriormente, pero están basadas en desarrollos que necesitaremos en el curso. En efecto, la prueba se basa en el siguiente hecho:

Proposición 2.1. *Sea B un movimiento browniano en ley. Para toda $t \geq 0$, si Δ_n es una sucesión de particiones de $[0, t]$ cuya norma tiende a cero, entonces la sucesión de variables aleatorias*

$$\sum_{t_i \in \Delta_n} \left(B_{t_i} - B_{t_{i-1}}\right)^2$$

converge en L_2 a t.

De hecho, para probar la proposición anterior, se verifica que

$$\sum_{t_i \in \Delta_n} \left[\left(B_{t_i} - B_{t_{i-1}}\right)^2 - (t_i - t_{i-1})\right]^2$$

converge a cero en L_2. Un movimiento browniano en ley $\tilde{B}$ es una martingala y por lo tanto le podemos asociar una modificación càdlàg B a partir del teorema de regularización de martingalas. (De revisar la demostración del teorema de regularización, se ve que no es necesario suponer que la filtración sea càd ó débilmente completa, lo que no se puede afirmar entonces es que B_t sea $\mathscr{F}_t$-medible.) Recordemos que una función càdlàg admite una cantidad a lo más numerable de discontinuidades. Por ser B una modificación de $\tilde{B}$, se ve que B también es un movimiento browniano en ley y por lo tanto, existe una sucesión de particiones Δ_n (y podemos suponer que las particiones se refinan) cuya norma tiende a cero tal que

$$\sum_{t_i \in \Delta_n} \left[\left(B_{t_i} - B_{t_{i-1}} \right)^2 - (t_i - t_{i-1}) \right]^2 \to 0$$

casi seguramente. Sin embargo, puesto que

$$\lim_{n \to \infty} \sum_{t_i \in \Delta_n} \left[\left(B_{t_i} - B_{t_{i-1}} \right)^2 - (t_i - t_{i-1}) \right]^2 \geq \sum_{s \leq t} \left(\Delta B_s \right)^2,$$

se ve que B no puede tener discontinuidades.

Prueba de la Proposición 2.1. Calculemos simplemente la norma $\mathscr{L}_2(\mathbb{P})$ al cuadrado de la diferencia entre las variables de interés:

$$\begin{aligned}\mathbb{E}\left(\left(\sum_i \left(B_{t_i} - B_{t_{i-1}} \right)^2 - t \right)^2 \right) &= \mathbb{E}\left(\left(\sum_i \left(\left(B_{t_i} - B_{t_{i-1}} \right)^2 - (t_i - t_{i-1}) \right) \right)^2 \right) \\ &= \mathbb{E}\left(\sum_i \left[\left(B_{t_i} - B_{t_{i-1}} \right)^2 - (t_i - t_{i-1}) \right]^2 \right),\end{aligned}$$

donde la última igualdad se justifica pues las variables

$$\left(\left(B_{t_i} - B_{t_{i-1}} \right)^2 - (t_i - t_{i-1}) \right)_i$$

son independientes y tienen media cero. Si ahora utilizamos el hecho de que $\mathbb{E}(X^4) = 3\sigma^2$ si X tiene distribución $N(0, \sigma^2)$, entonces

$$\begin{aligned}\sum_i \mathbb{E}\left(\left(\left(B_{t_i} - B_{t_{i-1}} \right)^2 - (t_i - t_{i-1}) \right)^2 \right) &= 2 \sum_i (t_i - t_{i-1})^2 \\ &\leq 2 \left| \Delta_n \right| t \to 0,\end{aligned}$$

lo cual demuestra la afirmación de la proposición. □

Como veremos, la proposición anterior es un caso particular de un resultado análogo para martingalas continuas que conforma uno de los pilares sobre los que se sostiene la teoría del cálculo estocástico. El teorema al que se hace referencia es el Teorema 1.3 del capítulo 4 de [**RY99**] y la demostración que se encuentra ahí utiliza solamente resultados básicos sobre martingalas como la desigualdad maximal de Doob. En [**KS91**] se encuentra el mismo teorema (teorema 5.1 del capítulo 1) para martingalas continuas de cuadrado integrable. En particular, se sigue que $\langle B \rangle : t \mapsto t$ es el único proceso creciente y adaptado a la filtración (aumentada) de B tal que $B^2 - \langle B \rangle$ es una martingala.

La Proposición 2.1 implica la imposibilidad de utilizar la teoría de integración de Lebesgue-Stieltjes para construir la integral estocástica, en efecto, recordemos del Teorema 1.1 que las sumas tipo Riemann-Stieltjes convergen para cualquier integrando continuo si y sólo si la función es de variación acotada. Sin embargo, la Proposición 2.1 nos ayudará a construir una teoría de integración basada en la transformada de martingala.

2. Ejemplo de una filtración que satisface las condiciones habituales

En esta sección se introduce un ejemplo (fundamental e importante) de una filtración que satisface las condiciones habituales. Sea $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad *completo* en el que está definido un movimiento browniano B. Sea $\mathscr{N}$ la clase de los conjuntos $\mathbb{P}$-nulos y definamos

$$\mathscr{F}_t = \sigma(B_s : s \leq t) \vee \mathscr{N}.$$

PROPOSICIÓN 2.2. *B es un $(\mathscr{F}_t)$-movimiento browniano y la filtración $(\mathscr{F}_t, t \geq 0)$ satisface las hipótesis habituales.*

DEMOSTRACIÓN. Notemos que $\mathscr{N}$ es independiente de cualquier σ-álgebra. Además, la independencia de los incrementos de B permite ver que si

$$A \in \sigma(B_s : s \leq t) \quad \text{y} \quad C \in \sigma(B_{t+s} - B_t : s \geq 0)$$

entonces A y C son independientes. Esto implica que si $N \in \mathscr{N}$

$$\mathbb{P}(A \cap C \cap N) = \mathbb{P}(A \cap C)\, \mathbb{P}(N) = \mathbb{P}(A)\, \mathbb{P}(C)\, \mathbb{P}(N) = \mathbb{P}(A \cap N)\, \mathbb{P}(B).$$

Por el lema de clases de Dynkin es entonces posible mostrar que C es independiente de $\mathscr{F}_t$ y que por lo tanto B es un $(\mathscr{F}_t)$-movimiento browniano.

Para cada $n \geq 1$ consideremos a la σ-álgebra

$$\mathscr{F}^n = \sigma\big(B_s - B_{t+1/n}, s \in [t + 1/n, t + 1]\big).$$

Puesto que B es un $(\mathscr{F}_t)$-movimiento browniano, se sigue que $\mathscr{F}^n$ es independiente de $\mathscr{F}_{t+1/n}$. Por lo tanto, si $A \in \mathscr{F}_{t+}$, $C \in \mathscr{F}_t$ y $D \in \mathscr{F}^n$, tenemos que

$$\mathbb{P}(A, C, D) = \mathbb{P}(A, C)\, \mathbb{P}(D) = \mathbb{E}(\mathbf{1}_C \mathbb{P}(\, A \mid \mathscr{F}_t))\, \mathbb{P}(D) = \mathbb{E}(\mathbf{1}_{C \cap D} \mathbb{P}(\, A \mid \mathscr{F}_t)).$$

Se concluye que

$$\mathbb{P}(\, A \mid \mathscr{F}_t, \mathscr{F}^n) = \mathbb{P}(\, A \mid \mathscr{F}_t).$$

Notemos que la sucesión de σ-álgebras $\mathscr{F}_t \vee \mathscr{F}^n, n \geq 1$ es creciente y que

$$\sigma\left(\bigcup_{n \geq 1} \mathscr{F}_t \vee \mathscr{F}^n\right) = \mathscr{F}_{t+1}.$$

En efecto, sólo debemos mostrar que si $s \in (t, t+1]$ entonces B_s es medible respecto de la σ-álgebra del lado izquierdo. Sin embargo, puesto que las trayectorias de B son continuas, vemos que $B_s = \lim_{n \to \infty} B_t + B_s - B_{t+1/n}$. Finalmente, utilizamos el teorema de convergencia de Lévy, mediante el que obtenemos la igualdad casi segura

$$\lim_{n \to \infty} \mathbb{P}(\, A \mid \mathscr{F}_t, \mathscr{F}^n) = \mathbb{P}(\, A \mid \mathscr{F}_1) = \mathbf{1}_A$$

pues $A \in \mathscr{F}_{t+} \subset \mathscr{F}_1$. Podemos entonces concluir que

$$\mathbf{1}_A = \mathbb{P}(A \mid \mathscr{F}_t)$$

casi seguramente. Sin embargo, como $\mathscr{F}_t$ es completa y $\mathbf{1}_A$ es igual casi seguramente a una variable $\mathscr{F}_t$ medible, vemos entonces que $\mathbf{1}_A$ es $\mathscr{F}_t$-medible, o en otras palabras, que $A \in \mathscr{F}_t$. □

3. La integral estocástica respecto del movimiento browniano

El objetivo de esta sección será construir una integral respecto de las trayectorias brownianas. Esto es, definir integrales del tipo

$$\int_0^t H_s \, dB_s.$$

Aquí, H sería una función continua, que podría ser aleatoria o no. El problema para definirla como una integral de Lebesgue-Stieltjes, aún cuando H sea determinista, es el siguiente.

Proposición 2.3. *Casi seguramente, las trayectorias de B son de variación infinita en $[u, v]$ para todo $u \leq v$.*

Demostración. Basta probar que para cada $u \leq v$ fijos, las trayectorias de B tienen variación infinita en $[u, v]$. Sin embargo, por la Proposición 2.1 aplicada al movimiento browniano $B_{\cdot+u} - B_u$, vemos que si Δ_n es una sucesión de particiones de $[u, v]$ cuya norma tiende a cero entonces

$$\sum_{\Delta_n} \left| B_{t_i} - B_{t_{i-1}} \right| \to \infty$$

en probabilidad. Así, al pasar a una subsucesión para que la convergencia sea casi segura, vemos que la variación de B en $[u, v]$ es casi seguramente infinita. □

Podríamos sin embargo, considerar a las sumas tipo Riemann

$$\sum_{\Delta_n} H_{t_i} \left(B_{t_i} - B_{t_{i-1}} \right)$$

y considerar el límite conforme $n \to \infty$. Sin embargo, puesto que las trayectorias de B son casi seguramente de variación infinita, el Teorema 1.1 implica que habrá funciones H para las cuáles no hay convergencia. La idea de Itô fue considerar un modo de convergencia distinto para las sumas tipo Riemann (la convergencia en probabilidad) además de restringir a los integrandos posibles. Por ejemplo, podemos inmediatamente integrar a B respecto de B (donde $H_t = B_t$): puesto que

$$\sum_{\Delta_n} B_{t_{i-1}} \left(B_{t_i} - B_{t_{i-1}} \right) = \frac{1}{2} \sum_{\Delta_n} \left(B_{t_i}^2 - B_{t_{i-1}}^2 \right) - \left(B_{t_i} - B_{t_{i-1}} \right)^2,$$

la Proposición 2.1 implica que

$$\sum_{\Delta_n} B_{t_{i-1}}\left(B_{t_i} - B_{t_{i-1}}\right) \to \frac{B_t^2 - t}{2}$$

en probabilidad. Aquí ya nos podemos dar cuenta de que necesariamente habrá una diferencia entre la integral de Lebesgue-Stieltjes y la integral respecto de las trayectorias brownianas, puesto que si f es continua y de variación localmente acotada y $f(0) = 0$, entonces

$$\int_0^t f\, df = \frac{f(t)^2}{2}$$

de acuerdo a la fórmula de integración por partes. La diferencia es por supuesto la existencia de la variación cuadrática para las trayectorias brownianas.

Sea Δ una $2n+1$-upla $t_0, \ldots, t_n, \tau_1, \ldots, \tau_n$ de reales en $[0, t]$, donde $0 = t_0 < t_1 < \cdots < t_n = t$, $0 \leq \tau_1, \ldots, \tau_{n-1} \leq t$ y $\tau_i \leq t_{i-1}$. (La idea es que $t_1, \ldots, t_n$ forman la partición de $[0, t]$ sobre la cual se consideran los incrementos del movimiento browniano y a las τ_i las utilizamos para evaluar al integrando.) A una tal $2n+1$-upla le llamaremos partición con instantes de evaluación. Definimos

$$d(\Delta) = \max_{1 \leq i \leq n-1} (t_i - \tau_i)\,.$$

Notemos que entonces

$$\tau_{i-1} \leq t_{i-1} \leq t_i \leq \tau_{i-1} + d(\Delta)\,.$$

Sea H un proceso estocástico adaptado a $\mathscr{F}_t^B, t \geq 0$ con trayectorias continuas. Definamos entonces

$$Y_t^{\Delta} = \sum_{\Delta} H_{\tau_i}\left(B_{t \wedge t_i} - B_{t \wedge t_{i-1}}\right).$$

Teorema 2.1. *Para toda $\varepsilon, \eta > 0$, existe $\delta > 0$ tal que si $d(\Delta), d(\Delta') < \delta$ entonces*

$$\mathbb{P}\Big(\left|Y_s^{\Delta} - Y_s^{\Delta'}\right| > \varepsilon\Big) < \eta.$$

La forma en que Itô escribe el teorema anterior es que Y^{Δ} converge en probabilidad conforme $d(\Delta) \to 0$.

Demostración. Sea Δ una sucesión con instantes de evaluación consistente de la partición $s_0, \ldots, s_m$ y de los instantes de evaluación $\sigma_1, \ldots \sigma_m$.

Si $t_0, \ldots, t_n$ es un refinamiento de $s_0, \ldots, s_m$ definamos

$$\tau_j = \sigma_i \quad \text{si} \quad (t_{j-1}, t_j) \subset (s_{i-1}, s_i)\,.$$

Definamos entonces Δ' como la partición con evaluadores conformada por $t_0, \ldots, t_n$ y $\tau_1, \ldots, \tau_n$, por lo que

$$Y^{\Delta} = Y^{\Delta'}.$$

Notemos que

$$d(\Delta') = \max t_{j+1} - \tau_j \leq \max s_{i+1} - \sigma_i = d(\Delta)\,,$$

por lo que Δ' es una especie de refinamiento de la partición con evaluadores.

Si Δ y $\tilde{\Delta}$ son dos particiones con evaluadores distintas, podemos entonces expresar a Y^{Δ} y a $Y^{\tilde{\Delta}}$ en términos de la misma partición, pero con instantes de evaluación distintos y escribir por tanto

$$Y^{\Delta} - Y^{\tilde{\Delta}} = \sum_i \left(H_{\tau_i} - H_{\tilde{\tau}_i}\right)\left(B_{t_i} - B_{t_{i+1}}\right).$$

Como H tiene trayectorias continuas, entonces para toda $\varepsilon, \eta > 0$ existe $\Delta > 0$ tal que

$$\mathbb{P}(|H_t - H_s| < \varepsilon \text{ si } |t-s| \leq \delta, s, t \leq 1) > 1 - \eta.$$

(Razón: se puede discretizar al evento anterior al considerar s, t racionales y utilizar la continuidad uniforme en el intervalo $[0,1]$.)

Sean $\varepsilon, \eta > 0$. Si $d(\Delta)\,, d\Big(\tilde{\Delta}\Big) < \delta$, donde $\delta = \delta(\varepsilon, \eta)$. Definamos

$$C_i = \left(H_{\tau_i} - H_{\tilde{\tau}_i}\right)\mathbf{1}_{\left|H_{\tau_i} - H_{\tilde{\tau}_i}\right| < \varepsilon},$$

por lo que

$$\mathbb{P}\Big(Y^{\Delta} - Y^{\tilde{\Delta}} \neq \sum C_i\left(B_{t_i} - B_{t_{i-1}}\right)\Big) < \eta.$$

Por otra parte

$$\mathbb{E}\bigg(\Big[\sum C_i\left(B_{t_i} - B_{t_{i-1}}\right)\Big]^2\bigg) = \sum \mathbb{E}\big(C_i^2\big)\left(t_i - t_{i-1}\right) \leq \varepsilon^2 T.$$

Por lo tanto

$$\mathbb{P}\left(\left|\sum_i C_i\left(B_{t_i} - B_{t_{i-1}}\right)\right| > \sqrt{\varepsilon}\right) \leq \varepsilon T$$

y por lo tanto

$$\mathbb{P}\Big(\Big|Y^{\Delta} - Y^{\tilde{\Delta}}\Big| > \sqrt{\varepsilon}\Big) \leq \varepsilon + \eta.$$

Es entonces sencillo ver (al considerar una sucesión de particiones evaluadoras Δ_n con $d(\Delta_n) \to 0$ y tales que Y^{Δ_n} converja en probabilidad) que existe una variable aleatoria Y tal que para toda $\varepsilon, \eta > 0$ existe $\delta > 0$ tal que si $d(\Delta) < \delta$ entonces

$$\mathbb{P}\Big(\Big|Y^{\Delta} - Y\Big| > \varepsilon\Big) < \eta. \qquad \square$$

Definición. La **integral estocástica** de H respecto de B en el intervalo $[0,1]$ es el límite en probabilidad de las sumas de Riemann

$$\sum_{\Delta} H_{\tau_i}\left(B_{t_i} - B_{t_{i-1}}\right)$$

(donde Δ es una partición evaluadora de $[0,1]$) conforme $d(\Delta) \to 0$. Se denotará por $\int_0^1 H_s\, dB_s$.

Claramente, los argumentos funcionan de igual manera en el intervalo $[u,v]$ en vez de $[0,1]$. Además, de acuerdo a la definición, tenemos que

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

Itô introdujo su célebre fórmula para calcular muchas otras integrales estocásticas.

El argumento anterior de hecho puede mejorarse para poder considerar una integral indefinida y probar propiedades de continuidad respecto del intervalo de integración. Antes de eso, enunciemos algunas propiedades de la integral estocástica.

PROPOSICIÓN 2.4. *Sean F y G dos procesos continuos y adaptados respecto de la filtración browniana. Entonces:*

(1) $\int_s^t dB_r = B_t - B_s$,
(2) *si* $\lambda, \mu \in \mathbb{R}$ *entonces* $\int_s^t \lambda F_s + \mu G_s\, dB_s = \lambda \int_s^t F_s\, dB_s + \mu \int_s^t G_s\, dB_s$ *y*
(3) *si* $r < s < t$ *entonces*

$$\int_r^s F_u\, dB_u + \int_s^t F_u\, dB_u = \int_r^t F_u\, dB_u.$$

EJERCICIO 2.1. Pruebe la proposición anterior.

PROPOSICIÓN 2.5. *Sea H un proceso adaptado con trayectorias continuas. Si existe una función continua tal que $\mathbb{E}\big(H_t^2\big) \leq M_t$ entonces*

$$\mathbb{E}\left(\left[\int_s^t H_r\, dB_r\right]^2\right) \leq \int_s^t M_r\, dr.$$

Sean $H^1, H^2, \ldots$ procesos estocásticos adaptados con trayectorias continuas y suponga que H^n converge a H uniformemente en probabilidad, es decir,

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{s\leq t} |H_s^n - H_s| > \varepsilon\right) = 0$$

para toda $\varepsilon > 0$. Entonces

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\int_0^t H_s^n\, dB_s - \int_0^t H_s\, dB_s\right| > \varepsilon\right) = 0.$$

DEMOSTRACIÓN. Se probará el resultado cuando $s = 0$. Sea (Δ_n) una sucesión cuyo paso tiende a cero y tal que la sucesión de integrales estocásticas elementales

$$I_n = \sum_{\Delta_n} H_{t_{i-1}} \left(B_{t_i} - B_{t_{i-1}}\right)$$

converge casi seguramente a la integral estocástica $\int_0^t H_r\, dB_r$. Entonces, por el lema de Fatou:

$$\mathbb{E}\left(\left[\int_0^t H_r\, dB_r\right]^2\right) \leq \liminf_{n\to\infty} \mathbb{E}\Big([I_n]^2\Big) \leq \liminf_{n\to\infty} \sum_n M_{t_{i-1}}\,(t_i - t_{i-1}) = \int_0^t M_r\, dr$$

puesto que M es Riemann integrable al ser continua.

Sea

$$C_s^n = \begin{cases} -\varepsilon & H_s - H_s^n \leq -\varepsilon \\ H_s - H_s^n & -\varepsilon < H_s - H_s^n < \varepsilon \\ \varepsilon & H_s - H_s^n \geq \varepsilon \end{cases}.$$

Al utilizar particiones aproximantes, es fácil ver que

$$\mathbb{P}\left(\int_0^t C_s^n\, dB_s \neq \int_0^t [H_s - H_s^n]\, dB_s\right) \leq \mathbb{P}\left(\sup_{s\leq t} |H_s - H_s^n| > \varepsilon\right).$$

Puesto que $C_s^n \leq \varepsilon$, se sigue que

$$\mathbb{P}\left(\left|\int_0^t C_s^n\, dB_s\right| \geq \sqrt{\varepsilon}\right) \leq \frac{1}{\varepsilon}\mathbb{E}\left(\left[\int_0^t C_s^n\, dB_s\right]^2\right) \leq \varepsilon t.$$

Por lo tanto

$$\mathbb{P}\left(\left|\int_0^t H_s\, B_s - \int_0^t H_s^n\, dB_s\right| > \varepsilon + \sqrt{\varepsilon}\right) \leq \mathbb{P}\left(\sup_{s\leq t} |H_s - H_s^n| > \varepsilon\right) + \varepsilon t$$

□

Para cada partición evaluadora definamos

$$Y_t^\Delta = \sum H_{\tau_i}\left(B_{t\wedge t_i} - B_{t\wedge t_{i-1}}\right).$$

Esta sería una versión discreta de la integral indefinida. Su principal característica es ser un proceso con trayectorias continuas y una martingala.

Proposición 2.6. *Para toda $\varepsilon, \eta > 0$, existe $\delta > 0$ tal que si $d(\Delta), d(\Delta') < \delta$ entonces*

$$\mathbb{P}\left(\sup_{s\leq t}\left|Y_s^\Delta - Y_s^{\Delta'}\right| > \varepsilon\right) < \eta.$$

Demostración. Sean $s_1, s_2, \ldots$ densos en $[0,1]$ y sean $t_1^m, \ldots, t_n^m$ los puntos obtenidos al refinar a Δ y a $\tilde{\Delta}$. Entonces, podremos escribir

$$Y_t^\Delta - Y_t^{\tilde{\Delta}} = \sum \left(H_{\tau_i} - H_{\tilde{\tau}_i}\right)\left(B_{t_i^m} - B_{t_{i-1}^m}\right).$$

Podemos seguir la prueba del Teorema 2.1 para deducir de la desigualdad de Doob (aplicada a la transformada de martingala de C por B) que

$$\mathbb{P}\left(\max_{1\leq i\leq n}\left|\sum_{i=1}^{k} C_i\left(B_{t_i^m}-B_{t_{i-1}^m}\right)\right|>\sqrt{\varepsilon}\right)\leq \varepsilon,$$

lo cual implicará que

$$\mathbb{P}\left(\sup_i\left|Y^{\Delta}_{t_I^m}-Y^{\tilde{\Delta}}_{t_i^m}\right|>\sqrt{\varepsilon}\right)\leq \varepsilon+\eta.$$

Al tomar el límite conforme $m\to\infty$ y utilizar la densidad de $\{s_1,s_2,\ldots\}$, obtenemos

$$\mathbb{P}\left(\sup_{s\leq 1}\left|Y^{\Delta}_{s}-Y^{\tilde{\Delta}}_{s}\right|>\sqrt{\varepsilon}\right)\leq \varepsilon+\eta,$$

lo cual nos permite concluir de manera análoga que en el Teorema 2.1. □

Corolario 1. *Existe un proceso estocástico con trayectorias continuas $H\cdot B$ tal que para todo $t\geq 0$*

$$\mathbb{P}\left((H\cdot B)_t=\int_0^t H_s\,dB_s\right)=1.$$

Demostración. Al considerar $\varepsilon_n=1/2^n$, vemos que existe δ_n tal que

$$\mathbb{P}\left(\sup_{s\leq t}\left|Y^{\Delta}_{s}-Y^{\Delta'}_{s}\right|>\varepsilon_n\right)<2^{-n}$$

si $d(\Delta),d(\Delta')<\delta_n$. Escogemos entonces una sucesión de particiónes Δ_n que se vayan refinando y tales que $d(\Delta)<\delta_n$. Se sigue entonces que

$$\sum_n\mathbb{P}\left(\sup_{s\leq t}\left|Y^{\Delta_n}_{s}-Y^{\Delta_{n+1}}_{s}\right|>\varepsilon_n\right)<\infty$$

y vemos por el lema de Borel-Cantelli que casi seguramente existe N tal que para toda $n\geq N$ se tiene que

$$\sup_{s\leq t}\left|Y^{\Delta_n}_{s}-Y^{\Delta_{n+1}}_{s}\right|\leq\varepsilon_n.$$

Se sigue entonces que la sucesión $\left(Y^{\Delta_n},n\geq 1\right)$ es casi seguramente de Cauchy uniforme en $[0,t]$ y por lo tanto converge uniformemente a un proceso $Y=(Y_s,s\in[0,t])$ con trayectorias continuas.

Por unicidad casi segura del límite en probabilidad se sigue que $Y_s=\int_0^s H_r\,dB_r$ casi seguramente. □

Una propiedad importante de la integral estocástica respecto de su parámetro, que ya se ha utilizado implícitamente es su carácter de martingala.

Proposición 2.7. *Si H es adaptado, continuo y existe $M>0$ tal que $\mathbb{E}\left(H_t^2\right)\leq M$ entonces $H\cdot B$ es una martingala cuadrado integrable con trayectorias continuas.*

Demostración. Notemos que $H \cdot B$ es límite de una sucesión de procesos adaptados a la filtración completada $(\mathscr{F}_t, t \geq 0)$. Puesto que cada $\mathscr{F}_t$ es completa, se sigue que $H \cdot B$ es adaptado.

Sea Δ_n es una sucesión de particiones de $[s,t]$ tal que la sucesión

$$I_n = \sum_{\Delta_n} H_{t_{i-1}} \left(B_{t_i} - B_{t_{i-1}}\right)$$

converge casi seguramente a $\int_s^t H_s \, dB_s$. Notemos que al condicionar (usando el hecho de que H es adaptado) se obtiene

$$\mathbb{E}\big(I_n^2\big) = \sum_{\Delta_n} \mathbb{E}\Big(H_{t_{i-1}}^2\Big) \left(t_i - t_{i-1}\right) \leq M\left(t-s\right).$$

Por lo tanto, (I_n) es una sucesión acotada en L_2 que converge casi seguramente a $\int_s^t H_r \, dB_r$. Por el lema de Fatou se sigue que

$$\mathbb{E}\left(\left[\int_s^t H_r \, dB_r\right]^2\right) \leq M\left(t-s\right) < \infty$$

y que

$$\mathbb{E}\left(\int_s^t H_r \, dB_r \,\middle|\, \mathscr{F}_s\right) = \lim_{n\to\infty} \mathbb{E}\left(I_n \,|\, \mathscr{F}_s\right) = 0.$$

Por lo tanto $H \cdot B$ es una martingala cuadrado integrable. $\square$

Pasemos ahora a la fórmula de Itô, que nos da una clase muy grande y útil de ejemplos de integrales estocásticas. Se trata de la versión estocástica del teorema de cambio de variable.

Teorema 2.2. *Sea $f : \mathbb{R} \to \mathbb{R}$ una función de clase C_2. Entonces*

$$f(B_t) = f(B_0) + \int_0^t f'(B_s)\, dB_s + \frac{1}{2}\int_0^t f''(B_s)\, ds.$$

La heurística de la prueba es sencilla: para cualquier partición $0 = t_0 < \cdots < t_n = t$ suficientemente fina, al realizar una expansión de Taylor de orden 2 se obtiene

$$f(B_{t_i}) - f\left(B_{t_{i-1}}\right) \approx f'\left(B_{t_{i-1}}\right)\left(B_{t_i} - B_{t_{i-1}}\right) + \frac{1}{2} f''\left(B_{t_{i-1}}\right)\left(B_{t_i} - B_{t_{i-1}}\right)^2.$$

Al sumar sobre i se obtiene

$$f(B_t) - f(B_{t_0}) = \sum_i f'\left(B_{t_{i-1}}\right)\left(B_{t_i} - B_{t_{i-1}}\right) + \sum_i \frac{1}{2} f''\left(B_{t_{i-1}}\right)\left(B_{t_i} - B_{t_{i-1}}\right)^2.$$

El primer sumando del lado derecho converge a la integral estocástica

$$\int_0^t f'(B_s)\, dB_s$$

mientras que el segundo debería converger hacia

$$\int_0^t \frac{1}{2} f''(B_s)\, ds.$$

Demostración. Trabajaremos en $[0,1]$.

Puesto que B tiene trayectorias continuas, el máximo de su valor absoluto en $[0,1]$ es finito casi seguramente y por eso, dada $\eta > 0$ existe $M > 0$ tal que

$$\mathbb{P}\left(\sup_{s \leq 1} |B_s| \geq M\right) < \eta.$$

Puesto que f'' es continua en $[-M, M]$, dada $\varepsilon > 0$ existe $\delta > 0$ tal que $|f''(y) - f''(x)| < \varepsilon$ si $x, y \in [-M, M]$ y $|y - x| < \delta$.

Puesto que las trayectorias de B son uniformemente continuas, existe $\gamma > 0$ tal que

$$\mathbb{P}\left(\sup_{\substack{|t-s|<\gamma \\ 0 \leq s,t \leq 1}} |B_t - B_s| \geq \delta\right) < \eta.$$

Por lo tanto, si

$$\Omega' = \left\{\sup_{s \leq 1} |B_s| < M, \sup_{\substack{|t-s|<\gamma \\ 0 \leq s,t \leq 1}} |B_t - B_s| < \delta\right\}$$

entonces $\mathbb{P}(\Omega') > 1 - 2\eta$.

Al realizar una expansión de Taylor de orden 2 se obtiene

$$f(B_t) - f(B_s) = f'(B_s)\,(B_t - B_s) + \frac{1}{2} f''(B_s)\,(B_t - B_s)^2 + \frac{1}{2} R_{s,t}\,(B_t - B_s)^2$$

para $0 \leq s \leq t \leq 1$, donde

$$R_{s,t} = f''(B_s + \theta\,(B_t - B_s)) - f''(B_s)$$

para alguna $\theta \in [0,1]$ (que es aleatoria).

Definamos ahora a los truncamientos

$$C_{s,t} = R_{s,t} \mathbf{1}_{|R_{s,t}| \leq \varepsilon}$$

y

$$E_t = f''(B_t)\,\mathbf{1}_{|f''(B_t)| \leq R}$$

donde $R = \max_{|x| \leq M} |f(x)|$. Notemos que en Ω', si $s < t < s + \gamma$ entonces

$$\begin{aligned} f(B_t) - f(B_s) = {} & f'(B_s)\,(B_t - B_s) + \frac{1}{2} f''(B_s)\,(t - s) \\ & + \frac{1}{2} C_{s,t}\,(B_t - B_s)^2 + \frac{1}{2} E_s \left[(B_t - B_s)^2 - (t - s)\right]. \end{aligned}$$

Si $0 = t_0 < \cdots < t_n = 1$ es cualquier partición de paso menor a γ, podemos por lo tanto escribir en Ω'

$$\begin{aligned} f(B_1) - f(0) = &\sum_i f'\left(B_{t_{i-1}}\right)\left(B_{t_i} - B_{t_{i-1}}\right) + \frac{1}{2} f''\left(B_{t_{i-1}}\right)(t_i - t_{i-1}) \\ &+ \sum_i \frac{1}{2} C_{t_i, t_{i-1}} \left(B_{t_i} - B_{t_{i-1}}\right)^2 \\ &+ \frac{1}{2} E_{t_{i-1}} \left[\left(B_{t_i} - B_{t_{i-1}}\right)^2 - (t_i - t_{i-1})\right]. \end{aligned}$$

Escogamos ahora la norma de la partición de tal manera que

$$\mathbb{P}\left(\left|\sum_i f'\left(B_{t_{i-1}}\right)\left(B_{t_i} - B_{t_{i-1}}\right) - \int_0^1 f'(B_s)\, dB_s\right| > \varepsilon\right) < \eta$$

y

$$\mathbb{P}\left(\left|\sum_i \frac{1}{2} f''\left(B_{t_{i-1}}\right)(t_i - t_{i-1}) - \int_0^1 f''(B_s)\, ds\right| > \varepsilon\right) < \eta.$$

Por otra parte, por definición de C_t vemos que

$$\mathbb{E}\left(\left|\sum_i \frac{1}{2} C_{t_{i-1}, t_i} \left(B_{t_i} - B_{t_{i-1}}\right)^2\right|\right) \leq \varepsilon.$$

Si además imponemos que $|t_i - t_{i-1}| < \varepsilon / R^2$ entonces

$$\begin{aligned} &\mathbb{E}\left(\left|\sum_i \frac{1}{2} E_{t_{i-1}} \left[\left(B_{t_i} - B_{t_{i-1}}\right)^2 - (t_i - t_{i-1})\right]\right|^2\right) \\ &\leq \frac{R^2}{4} \sum_i 2\,(t_i - t_{i-1})^2 \leq \frac{\varepsilon}{2} \leq \varepsilon. \end{aligned}$$

Así, podemos concluir que

$$\mathbb{P}\left(\left|\sum_i \frac{1}{2} C_{t_{i-1}, t_i} \left(B_{t_i} - B_{t_{i-1}}\right)^2\right| > \sqrt{\varepsilon}\right) < \sqrt{\varepsilon}$$

y

$$\mathbb{P}\left(\left|\sum_i \frac{1}{2} E_{t_{i-1}} \left[\left(B_{t_i} - B_{t_{i-1}}\right)^2 - (t_i - t_{i-1})\right]\right| > \sqrt{\varepsilon}\right) \leq \varepsilon.$$

Por lo tanto, si $\varepsilon < 1$:

$$\mathbb{P}\left(\left|f(B_t) - f(B_0) - \int_0^t f'(B_s)\, dB_s - \frac{1}{2}\int_0^t f''(B_s)\, ds\right| > 3\sqrt{\varepsilon}\right) \leq 2\eta + 2\varepsilon. \qquad \square$$

4. Ecuaciones diferenciales estocásticas conducidas por el movimiento browniano

Equipados con una noción de integral estocástica y conscientes de su diferencia fundamental con la integral de Lebesgue-Stieltjes, podemos analizar el concepto de ecuación diferencial estocástica. Cabe mencionar que ésta era la motivación original del [**Itô87**] para introducir la integral estocástica pues estaba interesado en verificar su intuición de que las soluciones a ecuaciones diferenciales estocásticas deberían ser procesos de Markov.

En este capítulo, nos interesaremos principalmente en ecuaciones del tipo

$$dX_t = \sigma(t, X_t)\ dB_t + b(t, X_t)\ dt.$$

La interpretación es que buscamos un proceso X con trayectorias continuas y adaptado a la filtración de B tal que para toda $t \geq 0$ se tenga que

$$X_t = x + \int_0^t \sigma(s, X_s)\ dB_s + \int_0^t b(s, X_s)\ ds. \tag{1}$$

Cuando σ es idénticamente igual a cero, nos reducimos al caso de las ecuaciones diferenciales ordinarias. Así puesto, la primera tarea que tendremos será la de estudiar existencia y unicidad para tales ecuaciones.

Intuitivamente, al discretizar una ecuación diferencial estocástica al utilizar una partición $0 = t_0 < t_1 < \cdots$ obtenemos una relación de recurrencia del tipo:

$$X_{t_{i+1}} = X_{t_i} + \sigma(t_i, X_{t_i}) \left[B_{t_{i+1}} - B_{t_i}\right] + b(t_i) \left[t_{i+1} - t_i\right].$$

(Formalmente, estaríamos aplicando el método de Euler a la ecuación (1).) Un tal proceso se puede pensar como la adición del ruido $\sigma(t_i, X_{t_i}) \left[B_{t_{i+1}} - B_{t_i}\right]$ a la evolución determinista

$$X_{t_{i+1}} = X_{t_i} + b(t_i) \left[t_{i+1} - t_i\right].$$

Por esta razón, un tal proceso tiene la propiedad de Markov a tiempo discreto. Por supuesto, hay dos preguntas naturales. La primera es si la propiedad de Markov también se vale para las soluciones a (1) y la segunda es si, cuando la norma de la partición tiende a cero, la solución de la ecuación de recurrencia converge en algún sentido a la solución de (1). Esto es, la pregunta sería sobre la convergencia del método de Euler (en algún sentido).

Pensemos en la siguiente situación como motivación: recordemos que hemos interpretado a una martingala como el proceso de ganancias que tenemos al someter cierto capital en un juego de apuestas justo. Por ejemplo, pensemos que $B_{t+s} - B_s$ es la ganancia (o pérdida en caso de que su signo sea negativo) en el intervalo $[s, s+t]$, aunque pensamos que el juego se desarrolla a tiempo continuo. Si el gobierno nos cobra (continuamente) impuestos sobre nuestras ganancias a tasa λ_+ y, en caso de tener capital negativo pedimos dinero prestado que nos genera un interés infinitesimal de tasa λ_-, estaríamos tentados a escribir nuestro capital al tiempo t como la solución a la ecuación diferencial estocástica

$$X_t = x + B_t - \int_0^t \lambda_- \mathbf{1}_{X_s<0} X_s \, ds - \int_0^t \lambda_+ X_s \, ds.$$

Cuando $\lambda_+ = \lambda_- = -\lambda$ obtenemos al célebre proceso de Ornstein-Uhlenbeck, que es la solución de

$$X_t = x + B_t + \lambda \int_0^t X_s \, ds.$$

Otro ejemplo concreto de ecuación diferencial estocástica es el de la exponencial estocástica del movimiento browniano. Se trata de la solución a la ecuación diferencial estocástica (EDE)

$$X_t = x + \int_0^t X_s \, dB_s. \tag{2}$$

Ejercicio 2.2. Probar que si f es una función continua de variación acotada entonces g satisface

$$g(t) = x + \int_0^t g(s) \ f(ds)$$

si y sólo si

$$g(t) = xe^{f(t)}.$$

Sugerencia: utilice integración por partes para ver que ge^{-f} es constante.

El caso del browniano es un tanto distinto.

Ejercicio 2.3. Pruebe que $X_t = xe^{B_t - t/2}$ satisface la ecuación (2). *Sugerencia:* Aplique la fórmula de Itô para procesos del tipo $f(t, B_t)$. Note que para obtener una conclusión de unicidad como la del caso de variación finita, hace falta una fórmula de integración por partes para integrales estocásticas.

Una manera de resolver ecuaciones diferenciales ordinarias es mediante el método de Picard. Esta misma idea funciona para ecuaciones diferenciales estocásticas, como hizo notar Itô. Para simplificar el enunciado del teorema de existencia y unicidad, haremos un supuesto de Lipschitz global en los coeficientes σ y b de la ecuación (1).

Teorema 2.3. *Suponga que σ y b son continuas y que existen constantes K y $t \geq 0$ tal que para toda $x, y \in \mathbb{R}$ se tiene que*

$$|\sigma(t, y) - \sigma(t, x)| \leq K \, |y - x| \quad y \quad |b(t, y) - b(t, x)| \leq K \, |y - x| \, .$$

Entonces

(1) *Dada $x \in \mathbb{R}$ existe un proceso X continuo y adaptado a la filtración de B tal que*

$$X_t = x + \int_0^t \sigma(s, X_s)\, dB_s + \int_0^t b(s, X_s)\, ds.$$

(2) *Si $\tilde{X}$ es continuo, adaptado y*

$$\tilde{X}_t = x + \int_0^t \sigma\Big(s, \tilde{X}_s\Big)\, dB_s + \int_0^t b\Big(s, \tilde{X}_s\Big)\, ds.$$

entonces X y $\tilde{X}$ son indistinguibles.

La prueba utilizará fuertemente el siguiente resultado conocido como la desigualdad (o lema) de Gronwall.

Lema 2. *Sea $T > 0$ y $g : [0, T] \to [0, \infty)$ medible y acotada. Si existen constantes $A, B \geq 0$ tales que*

$$g(t) \leq A + B \int_0^t g(s)\, ds \text{ para toda } t \in [0, T]$$

entonces

$$g(t) \leq Ae^{Bt} \text{ para toda } t \in [0, T].$$

Note que en particular, si $A = 0$ entonces $g = 0$.

Demostración. Al iterar la desigualdad, vemos que

$$g(t) \leq A + B \int_0^t A + \int_0^{t_1} Bg(t_2)\, dt_2\, dt_1 = A + ABt + B \int_0^t g(t_2)\, (t - t_2)\, dt_2$$

y al continuar se obtiene recursivamente

$$g(t) \leq A + ABt + A\frac{B^2 t^2}{2} + \cdots + A\frac{B^n t^n}{n!} + B^{n+1} \int_0^t g(t_{n+1})\, \frac{(t - t_{n+1})^n}{n!}\, dt_{n+1}.$$

Puesto que g es acotada, vemos que la integral del lado derecho de la desigualdad anterior converge a cero conforme $n \to \infty$ y por lo tanto

$$g(t) \leq \sum_{n=0}^{\infty} A\frac{B^n t^n}{n!} = Ae^{Bt}.$$ □

Es ilustrativo hacer el argumento de existencia y unicidad en el caso determinístico.

Prueba del Teorema 2.3 si $\sigma = 0$. Como veremos en la prueba del caso general, podemos suponer que existe una constante K' tal que $|b(t, x)| \leq K' + K\, |x|$ (aunque sea en conjuntos acotados del parámetro temporal).

Para la existencia, sea $X_t^0 = x$ y para $n \geq 0$ definimos recursivamente

$$X_t^{n+1} = x + \int_0^t b(s, X_s^n)\, ds.$$

Se sigue entonces que (X^n) es una sucesión de funciones continuas. La prueba terminará si mostramos que convergen uniformemente en compactos a una función (necesariamente) continua X. En efecto, podemos entonces tomar el límite en la definición de X^{n+1} y concluir por el teorema de convergencia acotada que

$$X_t = x + \int_0^t b(s, X_s)\ ds.$$

Para mostrar que la sucesión (X^n) converge uniformemente en compactos, utilizamos la hipótesis de Lipschitz para concluir que

$$\sup_{s\leq t} \left|X_s^{n+1} - X_s^n\right| \leq \int_0^t K\left|X_s^n - X_s^{n-1}\right|\ ds$$

y puesto que

$$\left|X_t^1 - X_t^0\right| \leq \int_0^t K' + K\left|x\right|\ ds = K''t,$$

se sigue que

$$\sup_{s\leq t} \left|X_s^{n+1} - X_s^n\right| \leq K''\frac{t^n}{n!}.$$

Podemos concluir entonces que

$$\sum_n \sup_{s\leq t} \left|X_s^{n+1} - X_s^n\right| < \infty$$

lo cual implica que X^n converge uniformemente en $[0, t]$.

Para la unicidad, suponemos que X y $\tilde{X}$ son dos soluciones a la ecuación diferencial 1 con $\sigma = 0$. Al utilizar la hipótesis de Lipschitz se obtiene

$$\left|X_t - \tilde{X}_t\right| \leq \int_0^t K\left|X_s - \tilde{X}_s\right|\ ds,$$

lo cual implica, por el lema de Gronwall, que $\left|X_t - \tilde{X}_t\right| = 0$ para toda t. □

Prueba del Teorema 2.3. Trabajaremos en el intervalo fijo $[0, 1]$. Notemos que existe una constante K' tal que $|\sigma(s, y)| \leq K|y| + K'$ para $s \leq 1$ y $y \in \mathbb{R}$ (y análogamente para b). En efecto, puesto que σ es continua entonces $K' = \sup_{s\leq 1}|\sigma(s, 0)| < \infty$. La hipótesis Lipschitz global implica que

$$|\sigma(s, y)| \leq |\sigma(s, 0)| + K|y|.$$

Probemos primero la unicidad. Se asumirá primero que σ y b son acotadas, digamos por M. Sean X y $\tilde{X}$ dos procesos continuos y adaptados que satisfacen la ecuación (1). Sea

$$\tilde{g}(t) = \mathbb{E}\left(\sup_{s\leq t}\left|X_s - \tilde{X}_s\right|\right).$$

Entonces, al utilizar la cota para σ y B, así como la desigualdad de Doob y la desigualdad $(a+b)^2 \leq 2\left(a^2+b^2\right)$ obtenemos

$$\begin{aligned}\tilde{g}(t) &\leq 2\mathbb{E}\left(\sup_{s\leq t}\left[\int_0^s \sigma(r,X_r)-\sigma\left(r,\tilde{X}_r\right)\,dB_r\right]^2+\sup_{s\leq t}\left[\int_0^t b(r,X_r)-b\left(r,\tilde{X}_r\right)\right]^2\right)\\ &\leq 32M^2t+8M^2t^2\leq 40M^2t<\infty.\end{aligned}$$

Definamos ahora a

$$g(t)=\mathbb{E}\left(\left[X_t-\tilde{X}_t\right]^2\right).$$

Puesto que $X-\tilde{X}$ tiene trayectorias continuas y su supremo en $[0,t]$ tiene momento de orden dos finito, vemos que g es una función acotada y continua.

Por otra parte, al utilizar la hipótesis de Lipschitz global y la cota para el segundo momento de una integral estocástica implica que

$$g(t)\leq 2K^2(1+t)\int_0^t g(s)\,ds;$$

si en el intervalo $[0,1]$ acotamos a $t\mapsto(1+t)$ por 2, vemos que se puede aplicar el lema de Gronwall y concluir que $g(t)=0$. Así, hemos probado que para toda t, $X_t=\tilde{X}_t$ casi seguramente (esto es, que X es modificación de $\tilde{X}$). Sin embargo, como ambos procesos tienen trayectorias continuas, entonces son indistinguibles.

Cuando σ y b son sólo continuas y no acotadas y suponemos que hay dos procesos X y $\tilde{X}$ continuos y adaptados que satisfacen (1) entonces definimos

$$\Omega_K=\left\{\sup_{t\leq T}|X_t|\leq K,\sup_{t\leq T}\left|\tilde{X}_t\right|\leq K\right\}.$$

Se tiene que

$$\lim_{K\to\infty}\mathbb{P}(\Omega_K)\to 1.$$

Si M es una cota para b y σ en $[0,T]\times[-M,M]$, definamos

$$b_M(t,y)=\begin{cases}-M & b(t,y)\leq -M\\ b(t,y) & -M\leq b(t,y)\leq M\\ -M & b(t,y)\geq M\end{cases}$$

y analogamente se define a σ_M. Puesto que en Ω_K, X y $\tilde{X}$ satisfacen la ecuación diferencial estocástica con coeficientes acotados B_M y σ_M, entonces $X_t=\tilde{X}_t$ casi seguramente en Ω_K. Así:

$$\mathbb{P}\left(X_t\neq\tilde{X}_t\right)\leq 1-\mathbb{P}(\Omega_K)\to_{K\to\infty}0.$$

Para la existencia, definamos $X_t^0 = x$ para toda $t \geq 0$ y recursivamente al proceso adaptado y continuo

$$X_t^{n+1} = x + \int_0^t \sigma(s, X_s^n)\ dB_s + \int_0^t b(s, X_s^n)\ ds.$$

Primero probaremos que la sucesión $X_t^k, k \geq 0$ converge en L_2 uniformemente en t en el intervalo $[0, 1]$. Sea M una cota para σ y b en $[0, 1] \times \{x\}$. Entonces la desigualdad de Jensen implica

$$\mathbb{E}\left(\sup_{s \leq t} \left[\int_0^s b\left(r, X_r^0\right)\ dr\right]^2\right) \leq M^2 t$$

y por desigualdad L_2 de Doob aplicada a la integral estocástica (que es una martingala) y la cota para el segundo momento de la integral estocástica

$$\mathbb{E}\left(\sup_{s \leq t} \left[\int_0^s \sigma\left(r, X_r^0\right)\ dBr\right]^2\right) \leq 4M^2 t.$$

La desigualdad $(a+b)^2 \leq 2\left(a^2 + b^2\right)$ implica entonces que

$$\mathbb{E}\left(\sup_{s \leq t} \left[X_s^1 - X_s^0\right]^2\right) \leq 10M^2 t.$$

Un argumento análogo, que además utiliza la hipótesis de Lipschitz global, muestra que

$$\mathbb{E}\left(\sup_{s \leq t} \left[X_s^{n+1} - X_s^n\right]^2\right) \leq 10K^2 \int_0^t \mathbb{E}\left(\sup_{r \leq s} \left[X_r^n - X_r^{n-1}\right]^2\right)\ ds$$

por lo que inductivamente se verifica la desigualdad

$$\mathbb{E}\left(\sup_{s \leq t} \left[X_s^{n+1} - X_s^n\right]^2\right) \leq \left(10K^2\right)^n 10M^2 \frac{t^{n+1}}{(n+1)!}.$$

Puesto que

$$\sum_n \mathbb{E}\left(\sup_{s \leq 1} \left[X_s^{n+1} - X_s^n\right]^2\right)^{1/2} < \infty,$$

podemos concluir que X_t^n converge en L_2 a

$$X_t = x + \sum_{n=1}^{\infty} \left(X_t^n - X_t^{n-1}\right)$$

y que además

$$\mathbb{P}\left(\sup_{t \leq 1} |X_t^n - X_t| > \varepsilon\right) \leq \frac{1}{\varepsilon^2} \left[\sum_{k=n+1}^{\infty} \mathbb{E}\left(\left[\sup_{t \leq 1} |X_t^n - X_t|\right]^2\right)\right] \to_{n \to \infty} 0.$$

Por lo tanto, existe una subsucesión n_k tal que casi seguramente

$$\sup_{s\leq 1} |X_t^{n_k} - X_t| \to_{n\to\infty} 0$$

por lo que X tiene trayectorias continuas y es adaptado. Por la convergencia uniforme, vemos que

$$\lim_{k\to\infty} \sup_{t\leq 1} \left| \int_0^t b(s, X_s)\ ds - \int_0^t b(s, X_s^{n_k})\ ds \right| = 0$$

casi seguramente. Finalmente, puesto que $\sigma(t, X_t^{n_k})$ converge uniformemente en probabilidad hacia $\sigma(t, X_t^{n_k})$, se puede aplicar el teorema de convergencia de integrales estocásticas para concluir que

$$\lim_{n\to\infty} \int_0^t \sigma(s, X_s^{n_k})\ dB_s = \int_0^t \sigma(s, X_s)\ dB_s$$

en probabilidad (y al pasar a una nueva subsucesión podemos suponer que la convergencia es casi segura) y que por lo tanto X satisface la ecuación (1). □

En particular, el caso en que $\sigma = 0$ nos reduce a la ecuación diferencial ordinaria

$$X_t' = b(t, X_t)\,.$$

El teorema de existencia y unicidad aplica cuando bajo una condición de Lipschitz global sobre b. Hay un célebre teorema, debido a Peano que afirma que hay existencia local para la ecuación anterior con la sola hipótesis de que b sea continua. Respecto a la unicidad, y para darse una idea de lo que permite probar el esquema de iteración de Picard y lo que es verdad, veamos que si b no depende de la variable temporal y es continua, entonces hay unicidad con la sola hipótesis $b > 0$. En efecto, notemos que si f satisface $f' = b \circ f$ con b positiva, entonces f es (continua y) estrictamente creciente. Si $i = f^{-1}$, se puede calcular la derivada de i y concluir que

$$i'(x) = \frac{1}{f' \circ i(x)} = \frac{1}{b \circ f \circ i(x)} = \frac{1}{b(x)}.$$

Concluimos que si f y $\tilde{f}$ ambas satisfacen la ecuación que nos interesa, sus inversas tienen la misma derivada, por lo que son iguales y por lo tanto $f = \tilde{f}$. De hecho, probemos unicidad en un un contexto más general cuando $\sigma = 0$: si b es continua, estrictamente positiva y $t \mapsto b(t, x)$ es no-decreciente. Sean x^1 y x^2 dos funciones diferenciables con derivadas y^1 y y^2 que satisfagan $y^i = b\big(t, x_t^i\big)$. Sea $x_t^3 = x_{\alpha t}^2$ con $\alpha > 1$ y notemos que su derivada, denotada y^3 está dada por $y_t^3 = \alpha b\big(\alpha t, x_t^3\big)$. Sea

$$\tau = \inf\left\{t \geq 0 : x_t^1 > x_t^3\right\}.$$

Si τ fuera finito entonces, puesto que $x_\tau^3 = x_\tau^1$ por continuidad y definición de τ, vemos que

$$y_\tau^3 = \alpha b\big(\alpha\tau, x_\tau^3\big) > b\big(\tau, x_\tau^3\big) = b\big(\tau, x_\tau^1\big) = y_\tau^1.$$

Lo anterior implica que $x^3 > x_1$ en una vecindad derecha de τ contradiciendo la definición de τ. Vemos por lo tanto que $x^1 \leq x^3$ y, al considerar $\alpha \to 1$, vemos que $x^1 \leq x^2$. Al intercambiar los roles de x^1 y x^2 nos damos cuenta de que $x^1 = x^2$.

Continuaremos con la razón fundamental por la cual Itô introdujo a la integral estocástica y a las ecuaciones diferenciales estocásticas asociadas: la construcción de procesos de Markov. La idea es que cuando los coeficientes σ y b que conducen a una ecuación diferencial estocástica no dependen de la variable temporal entonces la solución es un proceso de Markov.

Comencemos con la definición de un proceso de Markov y, de hecho, mejor motivémosla en el caso del movimiento browniano. La idea es ver al movimiento browniano como un proceso de Markov y para esto, quisíeramos definir al browniano que comienza en cualquier $x \in \mathbb{R}$ y definir un análogo de la matriz de transición. El browniano que comienza en x se define como el proceso $x + B$ y el sentido de esta definición es que como B_s y $B^s = (B_{t+s} - B_s, t \geq 0)$ son independientes y B^s es un browniano, si queremos la distribución condicional de $B_{t+s}, t \geq 0$ dado que $B_s = x$, estará dada por la distribución de $x + B^s$ que es la de $x + B$. Por otra parte, recordemos que si X es cadena de Markov con matriz de transición P entonces $\mathbb{P}(X_{n+m} = j \mid X_m = i) = P^n_{i,j}$ y que esto caracteriza a la cadena. El problema es que para el movimiento browniano, aunque definamos a $\mathbb{P}(B_{t+s} = y \mid B_s = x)$ como $\mathbb{P}(B_t + x = y)$, esta probabilidad será cero. Por esta razón se define al núcleo de transición P_t de tal manera que para todo $x \in \mathbb{R}$, $P_t(x, \cdot)$ es la medida dada por

$$P_t(x, A) = \mathbb{P}(x + B_t \in A) .$$

En general, se define a un **núcleo de medidas de probabilidad en** $\mathbb{R}$ como una función $N : \mathbb{R} \times \mathscr{B}_{\mathbb{R}} \to [0, \infty)$ tal que

- para toda $x \in \mathbb{R}$ $N(x, \cdot)$ es una medida de probabilidad en $\mathscr{B}_{\mathbb{R}}$ y
- para toda $A \in \mathscr{B}_{\mathbb{R}}$, $N(\cdot, A)$ es una función medible.

En el caso browniano no hemos probado la segunda condición. Sin embargo, notemos que la medibilidad es cierta cuando $A = (-\infty, y]$ para toda $y \in \mathbb{R}$. El lema de clases de Dynkin nos permite entonces obtener la medibilidad deseada. Como en el caso de matrices de transición, a las que podemos pensar como núcleos de medidas de probabilidad en algún conjunto finito, podemos definir el producto de núcleos de medidas de probabilidad, que no será en general conmutativo. Si M y N son núcleos de probabilidad en $\mathbb{R}$, definimos al núcleo NM por medio de la fórmula

$$NM(x, A) = \int N(y, A) \; M(x, dy) .$$

En el caso del movimiento browniano, vemos que

$$P_t P_s(x, (-\infty, z]) = \int \mathbb{P}(y + B_t \leq z) \; \mathbb{P}(x + B_s \in dy)$$

y entonces vemos que $P_tP_s(x,\cdot)$ es la convolución de una distribución normal con media x y varianza s con una normal centrada de varianza t. Se obtiene por lo tanto una normal con media x y varianza $s+t$, que por supuesto es la medida de probabilidad $P_{t+s}(x,\cdot)$. Por lo tanto, se obtiene la igualdad

$$P_{t+s} = P_tP_s,$$

que podemos interpretar como una versión de las ecuaciones de Chapman-Kolmogorov. A $(P_t, t \geq 0)$ se le conoce como semigrupo de transición del movimiento browniano. Para definir a un proceso de Markov (homogéneo) con valores en $\mathbb{R}$ se hace algo similar.

Definición. Un semigrupo de transición en $\mathbb{R}$ es una colección de núcleos de transición $N = (N_t, t \geq 0)$ tal que $N_tN_s = N_{t+s}$.

Un proceso estocástico X definido en el espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ es un **proceso de Markov** con semigrupo de transición N si

$$\mathbb{P}\big(X_{t+s} \in A \,\big|\, \mathscr{F}_s^X\big) = N_t(X_s, A)\,.$$

Equvalentemente, si para toda función $f : \mathbb{R} \to \mathbb{R}$ medible y acotada

$$\mathbb{E}\big(f(X_{t+s}) \,\big|\, \mathscr{F}_s^X\big) = \int f(y)\ N_t(X_s, dy)\,.$$

La anterior definición se puede escribir de manera más compacta al definir el semigrupo de operadores de transición asociado a N. Primero, dado N_t y una función medible y acotada, podemos definir a la función acotada N_tf de la siguiente manera:

$$N_tf(x) = \int f(y)\ N_t(x, dy)\,.$$

Esta función será medible; para probarlo, sólo notamos que cuando $f = \mathbf{1}_A$, entonces N_tf es medible por definición de núcleo de medidas de probabilidad. Luego, se extiende el resultado a funciones simples. Finalmente se aproxima a cualquier función medible y acotada por una sucesión de funciones medibles y uniformemente acotadas y se aplica el teorema de convergencia dominada para concluir que N_tf es el límite de una sucesión de funciones medibles y por lo tanto medible. La definición de proceso de Markov se puede entonces escribir de la manera siguiente:

$$\mathbb{E}\big(f(X_{t+s}) \,\big|\, \mathscr{F}_s^X\big) = N_tf(X_s)\,.$$

Finalmente, podemos enunciar el teorema de Itô.

Teorema 2.4. *Bajo las hipóteis del Teorema 2.3, sea X_t^x la (única) solución a la ecuación diferencial estocástica*

$$X_t^x = x + \int_0^t \sigma(X_s^x)\ dB_s + \int_0^t b(X_s^x)\ ds.$$

Entonces, X^x es un proceso de Markov homogéneo.

Vale la pena contrastar con el caso determinista en el que $\sigma = 0$. En este caso, notemos que

$$X^x_{t+s} = X^x_s + \int_0^t b\big(X^x_{r+s}\big)\, dr,$$

por lo que obtenemos la igualdad

$$X^x_{t+s} = X^{X^x_t}_t .$$

Esto se puede escribir como una propiedad de flujo de la función $F : (t, x) \mapsto X^x_t$:

$$F(t+s, x) = F(t, F(s, x)) .$$

Algo similar ocurre en el caso estocástico; sin embargo, como las funciones F_t también dependen del azar, debemos también pensar en cuestiones de medibilidad.

Prueba del Teorema 2.4. Dada una función medible y acotada $f : \mathbb{R} \to \mathbb{R}$, definimos $P_t f(x) = \mathbb{E}(f(X^x_t))$. A continuación, probaremos que si $\mathscr{F}_t = \sigma(B_s : s \leq t)$ entonces

$$\mathbb{E}\big(f\big(X^x_{t+s}\big) \,\big|\, \mathscr{F}_s\big) = P_t f(X^x_s) . \tag{3}$$

Puesto que X^x es adaptado a $(\mathscr{F}_s)$, la ecuación (3) implica que X^x es un proceso de Markov con semigrupo (P_t).

Para probar (3), se requieren algunos preliminares. Sea $B^s_t = B_{t+s} - B_s$. Entonces B^s es un movimiento browniano independiente de $\mathscr{F}_s$. Sea por otra parte $\tilde{X}^x$ el único proceso continuo y adaptado a la filtración canónica de B^s que resuelve la ecuación diferencial estocástica

$$\tilde{X}^x_t = x + \int_0^t \sigma(r, X^x_r)\, dB^s_r + \int_0^t b\Big(r, \tilde{X}^x_s\Big)\, dr.$$

Entonces $\tilde{X}^x_t = F_t(x, \omega)$ donde F_t es $\mathscr{B}_{\mathbb{R}} \otimes \mathscr{F}^{B^s}_t$-medible; esto se puede ver a partir del procedimiento de aproximación dado en la prueba del Teorema 2.3. Consideremos al proceso $\overline{X}$ dado por

$$\overline{X}^x_t(\omega) = \begin{cases} X^x_t(\omega) & t < s \\ F_{t-s}(X^x_s(\omega), \omega) & t > s \end{cases} .$$

Entonces $\overline{X}^x$ satisface la misma ecuación diferencial que X^x, pues es fácil verificar la igualdad casi segura

$$\int_0^t \sigma\Big(r, \tilde{X}^x_r\Big)\, dB^s_r = \int_s^{t+s} \sigma\Big(r, \tilde{X}^x_{r-s}\Big)\, dB_r.$$

Por el Teorema 2.3 vemos que $\overline{X}^x_{t+s} = X^x_{t+s}$ casi seguramente. Finalmente, puesto que $\mathscr{F}_s$ es independiente de $\mathscr{F}^{B^s}_t$ y X^x_s es $\mathscr{F}_s$-medible, entonces

$$\mathbb{E}\big(f\big(X^x_{t+s}\big) \,\big|\, \mathscr{F}_s\big) = \mathbb{E}(f(F_t(X^x_s, \cdot)) \mid \mathscr{F}_s) = P_t f(X^x_s) . \qquad \square$$

Pasamos a un fenómeno con el que se debe tener cuidado al tratar de aproximar ya sea integrales estocásticas o ecuaciones diferenciales ordinarias al aproximar al browniano por un proceso de variación finita. Un ejemplo de esto sería el substituir al browniano por la sucesión de procesos gaussianso lineales por pedazos como lo hace Paul Lévy. El fenómeno que ilustraremos a continuación se conoce con el nombre de Wong-Zakai quienes lo introdujeron en [**WZ65**]. Supongamos que B^n es una sucesión de procesos estocásticos cuyas trayectorias tienen casi seguramente variación finita y $B_0^n = 0$. Entonces, por la regla de la cadena, se sigue que

$$(B_t^n)^2 = 2\int_0^t B_s^n \, dB_s^n.$$

Esto implica que

$$\lim_{n\to\infty} 2\int_0^t B_s^n \, dB_s^n = B_t^2 \neq \int_0^t B_s \, dB_s.$$

El siguiente teorema no es más que una elaboración de esta idea.

Teorema 2.5. *Sea $f : \mathbb{R} \to \mathbb{R}$ con derivada continua y B_n una sucesión de procesos con trayectorias de variación finita que comienzan en cero y convergen casi seguramente al movimiento browniano además de ser casi seguramente uniformemente acotados en compactos. Entonces*

$$\lim_{n\to\infty} \int_0^t f(B_n(s)) \; B_n(ds) = \int_0^t f(B_s) \; dB_s - \frac{1}{2}\int_0^t f'(B_s) \; ds.$$

De igual manera, tenemos la versión para ecuaciones diferenciales estocásticas.

Teorema 2.6. *Suponga que σ y b no dependen de la variable temporal, que σ es derivable y que tanto σ, b como σ' son globalmente Lipschitz. Suponga además que existe $\varepsilon > 0$ tal que $\sigma \geq \varepsilon$. Sea B_n una sucesión de procesos con trayectorias de variación acotada que comienzan en cero y convergen casi seguramente al movimiento browniano uniformemente en compactos. Sea X_n la única solución a la ecuacion diferencial ordinaria*

$$X_n(t) = x + \int_0^t \sigma(X_n(s)) \; B_n(ds) + \int_0^t b(X_n(s)) \; ds.$$

Entonces X_n converge casi seguramente y uniformemente en compactos al único proceso X que satisface la ecuación diferencial estocástica

$$dX_t = \sigma(X_t) \; dB_t + b(X_t) + \sigma'(X_t) \; dt.$$

CAPÍTULO 3

Integración estocástica respecto de semi-martingalas continuas

En el capítulo anterior desarrollamos la teoría de la integral estocástica respecto del movimiento browniano y estudiamos algunas de sus propiedades. Esta integral se define de manera muy natural como límite de sumas de Riemann y sin embargo, presenta tanto similitudes con la integral de Lebesgue-Stieltjes (como un teorema de continuidad) y diferencias importantes (por ejemplo en el teorema de cambio de variable). Sin embargo propiedades como la asociatividad o la fórmula de integración por partes no las hemos podido analizar en su forma estocástica pues nuestros integradores se encuentran reducidos al movimiento browniano. En un inicio, quisieramos integrar respecto de integrales estocásticas (brownianas); resulta ser que el espacio de integradores se puede agrandar aún más sin mayor problema al de martingalas continuas. La herramienta principal será construir una variación cuadrática para martingalas continuas. La existencia de la variación cuadrática se puede hacer de diversas maneras: se puede estudiar directamente, como en [**Kal02**, **RY99**, **RW00**], o a través de la descomposición de Doob-Meyer como en [**KS91**]. La última alternativa se basa en un resultado más general y más difícil de demostrar. Es por esto que verificaremos la existencia de la variación cuadrática directamente.

1. Martingalas continuas y su variación cuadrática

Bajo ciertas condiciones, hemos visto que una integral estocástica es una martingala y que tiene trayectorias continuas. En esta sección abordaremos el estudio general de las martingalas con trayectorias continuas definidas en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ con filtración $(\mathscr{F}_t, t \geq 0)$. Asumiremos que se satisfacen las condiciones habituales. El que $\mathscr{F}_t$ contenga a todos los conjuntos $\mathbb{P}$-nulos de $\mathscr{F}$ implica que cualquier límite casi seguro de procesos adaptados es adaptado. La continuidad por la derecha hace que los tiempos opcionales sean tiempos de paro.

Sorprendentemente, salvo las martingalas continuas triviales, todas tienen variación no acotada y esto imposibilita la definición de una integral estocástica tipo Lebesgue-Stieltjes.

PROPOSICIÓN 3.1. *Sea $(M_t, t \geq 0)$ una martingala continua con trayectorias de variación acotada en intervalos compactos casi seguramente. Entonces M tiene trayectorias constantes casi seguramente.*

DEMOSTRACIÓN. Al restar el valor inicial, podemos suponer que $M_0 = 0$.

Supongamos primero que la variación V_t de M en $[0,t]$ es casi seguramente menor o igual que la constante $K > 0$. Notemos que entonces la variable aleatoria

$$\sup_{\substack{|s_2 - s_1| \leq \delta \\ s_1, s_2 \leq t}} |M_{s_2} - M_{s_1}|$$

se encuentra acotada por K.

Consideremos particiones Δ de $[0,t]$ de norma menor o igual a δ. Entonces, puesto que los incrementos de una martingala no tienen correlación:

$$\begin{aligned}
\mathbb{E}\left(M_t^2\right) &= \mathbb{E}\left(\left[\sum_{\Delta} M_{t_i} - M_{t_{i-1}}\right]^2\right) \\
&= \mathbb{E}\left(\sum_{\Delta} \left[M_{t_i} - M_{t_{i-1}}\right]^2\right) \\
&\leq \mathbb{E}\left(V_t \sup_{|s_1 - s_2| \leq \delta} |M_{s_2} - M_{s_1}|\right).
\end{aligned}$$

Por el teorema de convergencia acotada y la continuidad de las trayectorias de M, vemos que el lado derecho tiende a cero conforme la norma de la partición tiende a cero. Por lo tanto $M_t = 0$ casi seguramente. Así, vemos que M es una modificación del proceso $t \mapsto 0$ y al tener ambos trayectorias continuas, entonces M y 0 son indistinguibles.

Cuando la variación de M es sólo finita casi seguramente y no acotada por una constante, entonces consideramos a los tiempos de paro

$$S_K = \inf\left\{t \geq 0 : V_t > K\right\}.$$

Al utilizar la conclusión del párrafo anterior, notamos que la martingala M^{S_K} tiene trayectorias constantes y puesto que $S_K \wedge t \to t$ conforme $K \to \infty$ (pues la variación de M en $[0,t]$ es finita casi seguramente) vemos que M tiene trayectorias constantes casi seguramente. □

Sin embargo, justo como en el caso browniano, las martingalas continuas tienen variación cuadrática finita. La idea de la prueba es considerar primero el caso de martingalas cuadrado integrables y luego, descomponer a la martingala sobre una partición Δ como sigue:

$$M_t^2 - M_0^2 = 2\sum_{\Delta} M_{t_{i-1} \wedge t}\left[M_{t_i \wedge t} - M_{t_{i-1} \wedge t}\right] + \sum_{\Delta}\left[M_{t_i} - M_{t_{i-1}}\right]^2.$$

Si I_t^Δ denota al primer sumando del lado derecho y T_t^Δ al segundo, notemos que I^Δ es automáticamente una martingala. Por cálculos directos, mostraremos que para cada t fija, I_t^Δ converge en L_2 conforme $|\Delta| \to 0$ y por la desigualdad de Doob, obtendremos que la convergencia es uniforme sobre compactos y que por lo tanto el límite es un proceso continuo y adaptado. Así, veremos que T^Δ converge uniformemente en compactos y su límite, denotado por $\langle M \rangle$ será un proceso continuo adaptado y no-decreciente al que llamaremos variación cuadrática de M. Lo caracterizaremos como el único proceso tal que $M^2 - \langle M \rangle$ es una martingala.

Teorema 3.1. *Si M es una martingala continua y acotada, entonces existe un único proceso creciente, continuo, adaptado y nulo en cero, denotado por $\langle M \rangle$, tal que $M^2 - \langle M \rangle$ es una martingala. Además, para cualquier sucesión de particiones Δ_n cuyo paso tienda a cero, se tiene la convergencia en probabilidad*

$$\mathbb{P} - \lim_{n\to\infty} \sum_{\Delta_n} \left[M_{t_i} - M_{t_{i-1}}\right]^2 = \langle M \rangle.$$

Demostración. Supondremos, sin pérdida de generalidad, que $M_0 = 0$.

Para mostrar la unicidad, notamos que si A y B son dos procesos crecientes, nulos en cero, continuos y adaptados tales que $M^2 - A$ y $M^2 - B$ es una martingala entonces $B - A$ es una martingala con trayectorias continuas y de variación finita y nula en cero. Acabamos de mostrar que entonces $B - A$ tiene trayectorias constantes y como comienza en cero, vemos que $A = B$.

Probaremos ahora la existencia y primeras propiedades de la variación cuadrática. Sea $\Delta = \{0 = t_0 < t_2 < \cdots \leq t_m = t\}$ una partición de $[0, t]$ y consideremos a un refinamiento Δ' de Δ dado por $\Delta' = \{0 = t'_0 < t'_1 < \cdots < t'_n = t\}$. Así, cada intervalo $[t'_{j-1}, t'_j]$ determinado por Δ' está contenido dentro de uno determinado por Δ, digamos $[t_{i-1}, t_i]$. Podemos entonces expresar a la martingala I^Δ como una transformada de martingala asociada a la partición Δ' al definir $\tau_j = t_{i-1}$ si $[t'_{j-1}, t'_j] \subset [t_{i-1}, t_i]$ y entonces, notemos que

$$I_t^\Delta = \sum_j M_{\tau_j} \left[M_{t_{j-1}} - M_{t_j}\right].$$

Consideremos ahora dos particiones Δ y Δ' de $[0, t]$, donde no necesariamente alguna refina a la otra. Sin embargo, $\Delta \cup \Delta' = \{0 = t_0 < t_1 < \cdots\}$ es una nueva partición que refina a ambas y entonces vemos que existen $\tau_j, \tau'_j \leq t_{j-1}$ tales que

$$I_t^\Delta - I_t^{\Delta'} = \sum_i \left[M_{\tau_j} - M_{\tau'_j}\right] \left[M_{t_{j-1}} - M_{t_j}\right].$$

Escribamos $H_j = M_{\tau_j} - M_{\tau'_j}$, vemos que H_j es $\mathscr{F}_{t_{j-1}}$-medible, por lo que al utilizar la ortogonalidad de los incrementos de las martingalas se obtiene:

$$\mathbb{E}\left(\left[I_t^\Delta - I_t^{\Delta'}\right]^2\right) = \mathbb{E}\left(\sum_j H_j^2 \left[M_{t_{j-1}} - M_{t_j}\right]^2\right).$$

Si el máximo de las normas de las particiones $|\Delta|\vee|\Delta'|$ está acotado por δ entonces

$$\max_{j:\tau_j\leq t} H_j^2 \leq \max_{\substack{|s_2-s_1|\leq\delta\\ s_1,s_2\leq t}} |M_{s_2}-M_{s_1}|^{\,2}.$$

Denotemos por $C_{\delta,t}$ a la cota del lado derecho y notemos que, al ser M continua y acotada (digamos por K), $\lim_{\delta\to 0} C_{\delta_t}\to 0$ y $C_{\delta_t}\leq 4K^2$. Con estas definiciones se obtiene la cota

$$\mathbb{E}\left(\left[I_t^{\Delta}-I_t^{\Delta'}\right]^2\right)\leq \mathbb{E}\left(C_{\delta,t}T_t^{\Delta\cup\Delta'}\right)\leq \mathbb{E}\left(C_{\delta,T}^2\right)^{1/2}\mathbb{E}\left(\left[T_t^{\Delta\cup\Delta'}\right]^2\right)^{1/2}.$$

Por el teorema de convergencia acotada, vemos que $\mathbb{E}\left(C_{\delta,T}^2\right)^{1/2}\to 0$ conforme $\delta\to 0$. Mostraremos ahora que

$$\sup_{\Delta,\Delta'}\mathbb{E}\left(\left[T_t^{\Delta\cup\Delta'}\right]^2\right)<\infty,$$

En efecto, notemos que

$$\mathbb{E}\left(\left[T_t^{\Delta}\right]^2\right)=\mathbb{E}\left(\sum_i\left[M_{t_i}-M_{t_{i-1}}\right]^4+2\sum_{i<j}\left[M_{t_i}-M_{t_{i-1}}\right]^2\left[M_{t_j}-M_{t_{j-1}}\right]^2\right).$$

Puesto que M tiene trayectorias acotadas por K entonces

$$\mathbb{E}\left(\sum_i\left[M_{t_i}-M_{t_{i-1}}\right]^4\right)\leq 4K^2\mathbb{E}\left(\sum_i\left[M_{t_i}-M_{t_{i-1}}\right]^2\right)=4K^2\mathbb{E}\left(M_t^2\right).$$

Por otra parte, al condicionar por $\mathscr{F}_{t_i}$ y utilizar la ortogonalidad de los incrementos de una martingala vemos que

$$\begin{aligned}
&\mathbb{E}\left(2\sum_{i<j}\left[M_{t_i}-M_{t_{i-1}}\right]^2\left[M_{t_j}-M_{t_{j-1}}\right]^2\right)\\
&=\mathbb{E}\left(2\sum_{i<j}\left[M_{t_i}-M_{t_{i-1}}\right]^2\left[M_t-M_{t_i}\right]^2\right)\\
&\leq 8K^2\mathbb{E}\left(T_t^{\Delta}\right)\\
&=8K^2\mathbb{E}\left(M_t^2\right).
\end{aligned}$$

Se concluye entonces que

$$\mathbb{E}\left(\left[T_t^{\Delta\cup\Delta'}\right]^2\right)\leq 12K^2\mathbb{E}\left(M_t^2\right)\leq 12K^4$$

y por lo tanto

$$\lim_{|\Delta|\vee|\Delta'|\to 0}\mathbb{E}\left(\left[I_t^{\Delta}-I_t^{\Delta'}\right]^2\right)=0.$$

Por la desigualdad L_2 de Doob, obtenemos

$$\lim_{\delta\to 0}\sup_{|\Delta|,|\Delta'|\leq\delta}\mathbb{E}\left(\sup_{s\leq t}\left|I_s^{\Delta}-I_s^{\Delta'}\right|^2\right)=0.$$

Recursivamente se puede entonces construir una sucesión de particiones (Δ_n) que se refinen y tal que

$$\sum_n\mathbb{E}\left(\sup_{s\leq t}\left|I_s^{\Delta_{n+1}}-I_s^{\Delta_n}\right|^2\right)^{1/2}<\infty.$$

Vemos entonces que la serie es convergente si consideramos el momento de orden uno y por lo tanto

$$\sum_n\sup_{s\leq t}\left|I_s^{\Delta_{n+1}}-I_s^{\Delta_n}\right|<\infty$$

casi seguramente por lo que para cada $s\leq t$, la sucesión $I_s^{\Delta_n}$ converge casi seguramente a

$$I_s=I_s^{\Delta_0}+\sum_{n=0}^{\infty}I_s^{\Delta_{n+1}}-I_s^{\Delta_n}.$$

Se cumple además la desigualdad

$$\sup_{s\leq t}\left|I_s-I_s^{\Delta_m}\right|\leq\sum_{n\geq m}\sup_{s\leq t}\left|I_s^{\Delta_{n+1}}-I_s^{\Delta_n}\right|,$$

por lo que I^{Δ_n} converge a I uniformemente en $[0,t]$ y por lo tanto I es continuo. Al ser límite de procesos adaptados respecto de una filtración que satisface las condiciones usuales se sigue que I es adaptado. Se deduce por lo tanto que T^{Δ} converge uniformemente en $[0,t]$ al proceso continuo y adaptado $\langle M\rangle=M^2-I$. Para ver que $\langle M\rangle$ es creciente, consideramos a $t_1<t_2$ en el conjunto denso $\cup_n\Delta_n$. En cuanto Δ_n contenga tanto a t_1 como a t_2 se tendrá que $T_{t_1}^{\Delta_n}\leq T_{t_2}^{\Delta_n}$, lo cual explica por qué $\langle M\rangle$ es creciente en $\cup_n\Delta_n$. Por continuidad, vemos que $\langle M\rangle$ es no-decreciente.

Veamos ahora que $M^2-\langle M\rangle$ es una martingala. En efecto, notemos que $M^2-T^{\Delta_n}=I^{\Delta_n}$ es una martingala y que

$$\lim_{m\to\infty}\mathbb{E}\left(\left|I_t-I_t^{\Delta_m}\right|^2\right)\leq\lim_{m\to\infty}\sum_{n\geq m}\mathbb{E}\left(\left|I_t^{\Delta_{m+1}}-I_t^{\Delta_n}\right|^2\right)^{1/2}=0,$$

por lo que $I_t^{\Delta_n}$ converge en L_2 a I_t y por lo tanto también en L_1 (además de casi seguramente como ya habíamos demostrado). Esto nos dice que para cualquier $s\leq t$:

$$I_t=\lim_{n\to\infty}I_s^{\Delta_n}=\lim_{n\to\infty}\mathbb{E}\left(I_t^{\Delta_n}\,\middle|\,\mathscr{F}_s\right)=\mathbb{E}(I_t\mid\mathscr{F}_s).$$

Finalmente, si Δ'_n es cualquier sucesión de particiones tal que $|\Delta'_n| \to 0$, entonces

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\left(\sup_{s\leq t}\left|I_s^{\Delta'_n} - I_s\right| > \varepsilon\right) &\leq \lim_{n\to\infty} \mathbb{P}\left(\sup_{s\leq t}\left|I_s^{\Delta'_n} - I_s^{\Delta_n}\right| > \varepsilon/2\right) \\ &\quad + \mathbb{P}\left(\sup_{s\leq t}\left|I_s^{\Delta'_n} - I_s^{\Delta_n}\right| > \varepsilon/2\right) \\ &= 0\end{aligned}$$

puesto que

$$\mathbb{P}\left(\sup_{s\leq t}\left|I_s^{\Delta'_n} - I_s^{\Delta_n}\right| > \varepsilon/2\right) \leq \frac{2}{\varepsilon}\mathbb{E}\left(\sup_{s\leq t}\left|I_s^{\Delta'_n} - I_s^{\Delta_n}\right|\right) \to 0.$$

Por lo tanto, $T^{\Delta'_n}$ converge a $\langle M\rangle$ uniformemente en $[0,t]$ en probabilidad. □

La demostración anterior es una simplificación de la encontrada en la prueba del Teorema IV.1.3 de [**RY99**]. La simplificación se basa en la idea de Itô de comparar dos sumas de Riemman respecto de particiones distintas al considerar su unión. Como veremos en este capítulo, el teorema anterior abre la puerta hacia una teoría de la integración estocástica para martingalas continuas. En [**RW00**], también podemos encontrar una prueba simple de un resultado similar al teorema anterior que es suficiente para construir la integral estocástica (cf. el Teorema 30.1 en la página 53). Finalmente, la prueba de la existencia de la variación cuadrática es mucho más simple en el caso de trayectorias continuas que en el caso de trayectorias càdlàg. La última se basa en la célebre descomposición de Doob-Meyer (propuesta por Doob y probada por Meyer en [**Mey62**] y [**Mey63**]) cuya demostración inicial requería herramientas muy profundas de la teoría general de procesos estocásticos. El propio Meyer se lamentaba de que para poder enseñar cálculo estocástico respecto de semimartingalas càdlàg *se requiere un curso de un semestre... nada más para las definiciones.* Cabe mencionar que recientemente se ha publicado una prueba más sencilla de la descomposición de Doob-Meyer en [**BSV12**]. Tal vez esta prueba pueda reducir el tiempo necesario para un curso de cálculo estocástico respecto de semimartingalas càdlàg. Mientras eso se decide, seguiremos con la presentación del caso de las (semi)martingalas continuas, que es al que más aplicaciones se le han dado.

El Teorema 3.1 podría parecer limitado pues, al imponer que la martingala sea acotada, deja fuera incluso al movimiento browniano. Una sencilla técnica, llamada de localización, nos permite extender el teorema anterior a las llamadas martingalas locales y a las semimartingalas. Recordemos que si T es un tiempo aleatorio y X es un proceso estocástico entonces X^T denota al proceso X detenido en T, dado por $X^T_t = X_{t\wedge T}$.

Proposición 3.2. *Si M es una martingala continua y acotada y T es un $(\mathscr{F}_t)$-tiempo de paro entonces $\langle M^T\rangle = \langle M\rangle^T$.*

Demostración. Si M es acotada, entonces M^T es también una martingala acotada (respecto a la misma filtración), por lo que tiene sentido cuestionarse sobre su variación cuadrática. Puesto que por definición $M^2 - \langle M \rangle$ es una martingala entonces $\left(M^2 - \langle M \rangle\right)^T = \left(M^T\right)^2 - \langle M \rangle^T$ es una martingala y, por unicidad de la variación cuadrática, $\langle M \rangle^T = \langle M^T \rangle$. □

Con la propiedad anterior podremos dar una primera extensión del Teorema 3.1 que cubra al movimiento browniano.

Definición. Una **martingala local continua** es un proceso estocástico $M = (M_t, t \geq 0)$ con trayectorias continuas tal que existe una sucesión de tiempos de paro $T_1 \leq T_2 \leq \cdots$ tales que $T_n \to \infty$ casi seguramente y

(1) M_0 es $\mathscr{F}_0$-medible y
(2) $(M - M_0)^{T_n}$ es una martingala acotada.

Si M es cualquier proceso estocástico con trayectorias continuas y T es un tiempo de paro tal que M_0 es $\mathscr{F}_0$-medible y $(M - M_0)^T$ es una martingala acotada, decimos que el tiempo de paro T **reduce** al proceso M. Por ejemplo, el movimiento browniano es una martingala local. De hecho, cualquier martingala con trayectorias continuas es una martingala local, como se puede ver al definir $T_n = \inf \{t \geq 0 : |M_t - M_0| \geq n\}$.

Corolario 2. *Si M es una martingala local, existe un único proceso creciente, nulo en cero, continuo y adaptado $\langle M \rangle$ tal que $M^2 - M$ es una martingala local. Además, para cualquier sucesión de particiones Δ_n sin puntos de acumulación cuyo paso tiende a cero, la sucesión T^{Δ_n} converge a $\langle M \rangle$ uniformemente en compactos en probabilidad.*

Demostración. Probemos la existencia; al substraer M_0, podemos suponer que $M_0 = 0$. Sea T_n una sucesión creciente de tiempos de paro tal que $T_n \to \infty$ casi seguramente y cada T_n reduce a M. La martingala acotada M^{T_n} admite una variación cuadrática $\langle M^{T_n} \rangle$. El corolario 2 nos dice que si $m \leq n$ entonces $\langle M^{T_m} \rangle = \langle M^{T_n} \rangle^{T_m}$. Por lo tanto, podemos definir al proceso estocástico $\langle M \rangle$ mediante

$$\langle M \rangle_t = \langle M^{T_n} \rangle \quad \text{si } t \leq T_n.$$

Puesto que $\langle M \rangle^{T_n} = \langle M^{T_n} \rangle$ por construcción, vemos que si

$$S_n = \inf \{t \geq 0 : \langle M \rangle \geq n\}$$

entonces $\left(M^2 - \langle M \rangle\right)^{T_n \wedge S_n} = \left(M^2\right)^{T_n \wedge S_n} - \langle M^{T_n \wedge S_n} \rangle$ es una martingala acotada. Por lo tanto $M^2 - \langle M \rangle$ es una martingala local continua.

Para la unicidad, supongamos que A y B son dos procesos crecientes, nulos en cero, continuos y adaptados tales que $M^2 - A$ y $M^2 - B$ son martingalas locales continuas. Entonces $B - A$ es una martingala local continua con trayectorias de variación acotada en compactos. Si (T_n) es una sucesión creciente de tiempos de paro que reducen a M y $S_n = \inf\{t \geq 0 : A_t \geq n \text{ ó } B_t \geq n\}$entonces $(B - A)^{T_n \wedge S_n}$ es una martingala acotada continua con trayectorias de variación finita. Por la Proposición 3.1, $(B - A)^{T_n \wedge S_n}$ tiene trayectorias constantes. Puesto que $T_n \wedge S_n \to \infty$ casi seguramente entonces

$$
\begin{aligned}
&\mathbb{P}(\exists t \geq 0, B_t \neq A_t) \\
&= \sum_n \mathbb{P}(\exists t \in [S_n \wedge T_n, S_{n+1} \wedge T_{n+1}), B_t \neq A_t) \\
&= \sum_n \mathbb{P}\Big(\exists t \in [S_n \wedge T_n, S_{n+1} \wedge T_{n+1}), B_t^{S_{n+1} \wedge T_{n+1}} \neq A_t^{S_{n+1} \wedge T_{n+1}}\Big) \\
&= 0.
\end{aligned}
$$

Finalmente, sea Δ_n una sucesión de particiones sin punto de acumulación cuyo paso tiende a cero. Sea $T_n, n \geq 1$ una sucesión creciente de tiempos de paro que crece a infinito y tal que cada T_n reduce a M. Sea además $S_n = \inf\{t \geq 0 : \langle M \rangle \geq n\}$. Consideremos a $\delta > 0$ y n tal que

$$\mathbb{P}(S_n \wedge T_n \leq t) < \delta.$$

Sea además m_0 tal que si $m \geq m_0$ entonces

$$\mathbb{P}\left(\sup_{s \leq t} \left|T^{\Delta_m}\left(M^{S_n \wedge T_n}\right) - \langle M^{S_n \wedge T_n} \rangle\right| > \varepsilon\right) \leq \delta.$$

Entonces

$$
\begin{aligned}
&\mathbb{P}\left(\sup_{s \leq t} \left|T^{\Delta_m} - \langle M \rangle\right| > \varepsilon\right) \\
&\leq \delta + \mathbb{P}\left(\sup_{s \leq t} \left|T^{\Delta_m}(M) - \langle M \rangle\right| > \varepsilon, t \leq S_n \wedge T_n\right) \\
&\leq 2\delta.
\end{aligned}
$$

□

Proposición 3.3. *Sea M una martingala local continua. Entonces los intervalos de constancia de M y de $\langle M \rangle$ coinciden.*

La extensión final del Teorema 3.1 que analizaremos es en el contexto de las semimartingalas continuas.

Definición. Una **semimartingala continua** es un proceso estocástico X con trayectorias continuas que se puede descomponer como $M + A$ donde M es una martingala local continua y A es un proceso de variación acotada en compactos.

Notemos que la descomposición es única si imponemos la condición $A_0 = 0$, gracias al Teorema 3.1

Corolario 3. *Si $X = M + A$ es una semimartingala continua, entonces $T^{\Delta}(X)$ converge uniformemente en compactos a $\langle M \rangle$ en probabilidad.*

DEMOSTRACIÓN. Si Δ es una partición sin puntos de acumulación entonces

$$\begin{aligned}
T_t^\Delta &= \sum_\Delta \left[X_{t_i\wedge t} - X_{t_{i-1}\wedge t}\right]^2 \\
&= \sum_\Delta \left[M_{t_i} - M_{t_{i-1}\wedge t}\right]^2 \\
&+ 2\sum_\Delta \left[M_{t_i\wedge t} - M_{t_{i-1}\wedge t}\right]\left[A_{t_i\wedge t} - A_{t_{i-1}\wedge t}\right] \\
&+ \sum_\Delta \left[A_{t_i\wedge t} - A_{t_{i-1}\wedge t}\right]^2 .
\end{aligned}$$

Puesto que A tiene variación finita en compactos entonces el tercer sumando converge a cero uniformemente en $[0,t]$ casi seguramente si $|\Delta| \to 0$. Por otra parte si $V_t(A)$ denota la variación de A en $[0,t]$ y la norma de la partición es menor a δ entonces

$$\left|\sum_\Delta \left[M_{t_i\wedge t} - M_{t_{i-1}\wedge t}\right]\left[A_{t_i\wedge t} - A_{t_{i-1}\wedge t}\right]\right| \leq \sup_{\substack{|s_2-s_1|\leq\delta \\ s_1,s_2\leq t}} |M_{s_2} - M_{s_1}|\, V_t(A),$$

donde el lado derecho converge a cero uniformemente en compactos casi seguramente conforme $\delta \to 0$ gracias a la continuidad de las trayectorias de M. □

Ahora pasaremos a la construcción de la covariación entre dos martingalas locales continuas. Para esto, recordemos la *fórmula de polarización* $(a+b)^2 - (a-b)^2 = 4ab$. Sean M y N dos martingalas locales continuas y definamos la **covariación** entre M y N, denotada $\langle M,N\rangle$ por medio de la fórmula

$$\langle M,N\rangle = \frac{\langle M+N\rangle - \langle M-N\rangle}{4}.$$

COROLARIO 4. *Sean M y N martingalas locales continuas. Entonces $\langle M,N\rangle$ es el único proceso continuo, nulo en cero, con trayectorias de variación finita y adaptado tal que $MN - \langle M,N\rangle$ es una martingala continua. Sea Δ_n es una sucesión de particiones de $[0,\infty)$ sin puntos de acumulación cuya norma tiende a cero. Entonces la sucesión de procesos*

$$T_t^{\Delta_n}(M,N) = \sum_{\Delta_n} \left(M_{t_i\wedge t} - M_{t_{i-1}\wedge t}\right)\left(N_{t_i\wedge t} - N_{t_{i-1}\wedge t}\right)$$

converge uniformemente en compactos en probabilidad a (M,N).

EJERCICIO 3.1. Sean M y N dos martingalas locales continuas e independientes. Calcular $\langle M,N\rangle$.

2. La integral estocástica respecto de martingalas continuas cuadrado integrables

Pasemos ahora al estudio de una clase intermedia entre las martingalas continuas y acotadas y las martingalas locales continuas. Esta clase será fundamental para la construcción de la integral estocástica. Consideremos al conjunto H_2 que consta de las martingalas continuas M acotadas en L_2, esto es, tales que

$$\sup_t \mathbb{E}\big(M_t^2\big) < \infty.$$

Con esta hipótesis, sabemos que M_t converge casi seguramente y en L_2 conforme $t \to \infty$ a una variable M_∞ y que además

$$M_t = \mathbb{E}(\, M_\infty \mid \mathscr{F}_t)\, .$$

Hemos mostrado que a cada $M \in H_2$ le podemos asociar la variable M_∞ en L_2, por lo que H_2 se podría pensar como encajado en el espacio de Hilbert L_2. De hecho, claramente H_2 es un espacio vectorial. Además, lo podemos dotar del producto interior dado por

$$(M, N) \mapsto \mathbb{E}(M_\infty N_\infty)\, ,$$

que nos genera la norma en H_2 dada por

$$\|M\|_{H_2} = \|M_\infty\|_2 = \lim_{t\to\infty} \|M_t\|_2.$$

Ahora utilizaremos el encaje de H_2 en L_2 para probar que con este producto interior, H_2 resulta ser un espacio de Hilbert. En efecto, si M^n es una sucesión de Cauchy en H_2, escojamos una subsucesión (n_k) tal que

$$\sum_k \|M_\infty^{n_{k+1}} - M_\infty^{n_k}\|_2 < \infty.$$

Entonces, por la desigualdad L_2 de Doob y la desigualdad de Jensen, vemos que

$$\sum_k \sup_s |M_s^{n_{k+1}} - M_s^{n_k}| < \infty$$

casi seguramente. Por lo tanto, M^{n_k} converge uniformemente de manera casi segura y a cada tiempo fijo en L_2 a un proceso M que necesariamente tiene trayectorias continuas. Por la convergencia en L_2, vemos que M es una martingala y por el lema de Fatou se sigue que $M \in H_2$. Al utilizar el carácter Cauchy de (M_∞^n, n), vemos que M^n converge a M en H_2. Por lo tanto H_2 es completo y así vemos que se trata de un espacio de Hilbert.

Continuemos con nuestra construcción de la integral estocástica al introducir un segundo espacio de Hilbert, el espacio de integrandos admisibles.

Definición. Sea $M \in H_2$. Se define al espacio $\mathscr{L}_2(M)$ como el conjunto de procesos estocásticos H que son progresivamente medibles y

$$\mathbb{E}\bigg(\int_0^\infty H_s^2 \, d\langle M\rangle_s\bigg) < \infty.$$

Notemos que si en $\mathbb{R}_+ \times \Omega$ definimos la medida

$$\mathbb{P}_M(A) = \mathbb{E}\left(\int \mathbf{1}_A(s,\omega)\, d\langle M\rangle_s(\omega)\right)$$

entonces $\mathscr{L}_2(M)$ es un subconjunto del espacio $\mathscr{L}_2(\mathbb{P}_M)$. Denotaremos por $L_2(M)$ al conjunto de clases de equivalencia de elementos de $\mathscr{L}_2(M)$ que son iguales $\mathbb{P}_M$-casi dondequiera. Así, $L_2(M)$ se convierte en un espacio de Hilbert si se utiliza el producto interior

$$(H,K) = \mathbb{E}\left(\int_0^\infty H_s^2 K_s^2\, d\langle M\rangle_s\right).$$

Hay dos maneras de construir la integral estocástica: la manera directa mediante un argumento de densidad y la manera abstracta basada en el teorema de representación (de funcionales lineales continuas en un espacio de Hilbert) de Riesz. Para ambas requerimos las desigualdades de Kunita-Watanabe que introducimos a continuación.

Teorema 3.2 (Desigualdades de Kunita-Watanabe, [**KW67**]). *Sean M y N dos martingalas locales continuas y H y K dos procesos medibles. Entonces*

$$\int_0^\infty |K_s H_s|\; d\left|\langle M,N\rangle_s\right| \leq \int_0^\infty H_s^2\, d\langle M\rangle_s \times \int_0^\infty K_s^2\, d\langle N\rangle_s.$$

Como consecuencia, vemos que si $H \in L_2(M)$ y $K \in L_2(N)$ entonces

$$\mathbb{E}\left(\int_0^\infty |K_s H_s|\; d\left|\langle M,N\rangle_s\right|\right) \leq \|H\|_{L_2(M)}\|K\|_{L_2(N)}.$$

Demostración. Utilicemos la notación $\langle M,N\rangle_s^t = \langle M,N\rangle_t - \langle M,N\rangle_s$ para $s \leq t$.

Notemos primero que casi seguramente, para cualquier pareja de racionales no-negativos s,t con $s < t$ se tiene que

$$\left(\langle M,N\rangle_s^t\right)^2 \leq \langle M\rangle_s^t \langle N\rangle_s^t.$$

En efecto, basta utilizar la desigualdad de Cauchy para verificar que

$$\left[\sum \left(M_{t_i} - M_{t_{i-1}}\right)\left(N_{t_i} - N_{t_{i-1}}\right)\right]^2 \leq \left[\sum \left(M_{t_i} - M_{t_{i-1}}\right)^2\right]\left[\sum \left(N_{t_i} - N_{t_{i-1}}\right)^2\right]$$

y tomar el límite sobre particiones en $[s,t]$ cuya norma tiende a cero.

Por lo anterior, vemos que

$$\int_s^t |d\langle M,N\rangle_r| \leq \sqrt{\langle M\rangle_s^t \langle N\rangle_s^t}.$$

En efecto, recordemos que la variación total de la medida $\langle M, N\rangle$ es la medida inducida por la covariación $\langle M, N\rangle$, misma que se puede obtener, en un intervalo $[s, t]$ como supremo de sumas de incrementos en valor absoluto. En otras palabras:

$$\int_s^t |d\langle M, N\rangle_r| = V(\langle M, N\rangle, [s, t]) = \sup_\Delta \sum \left|\langle M, N\rangle_{t_{i-1}}^{t_i}\right|$$

Por otra parte, al utilizar la desigualdad del párrafo anterior en cada uno de los sumandos seguido de la desigualdad de Cauchy de nuevo, vemos que

$$\begin{aligned}\int_s^t |d\langle M, N\rangle_r| &\leq \sup_\Delta \sum \left(\langle M\rangle_{t_{i-1}}^{t_i}\right)^{1/2} \left(\langle N\rangle_{t_{i-1}}^{t_i}\right)^{1/2} \\ &\leq \sup_\Delta \left[\sum \langle M\rangle_{t_{i-1}}^{t_i}\right]^2 \left[\sum \langle N\rangle_{t_{i-1}}^{t_i}\right]^2 \\ &= \left(\langle M\rangle_s^t\right)^{1/2} \left(\langle N\rangle_s^t\right)^{1/2}.\end{aligned}$$

Generalizaremos el párrafo anterior de la siguiente manera: para todo A boreliano de $\mathbb{R}_+$ se tiene la desigualdad

$$\int_A |d\langle M, N\rangle| \leq \left(\int_A d\langle M\rangle\right)^{1/2} \left(\int_A d\langle N\rangle\right)^{1/2}.$$

En efecto, si A es la unión ajena finita de los intervalos $(s_i, t_i]$, entonces se deduce de la desigualdad del párrafo anterior y la desigualdad de Cauchy que

$$\begin{aligned}\int_A |d\langle M, N\rangle| &= \sum_i \int_{s_i}^{t_i} |d\langle M, N\rangle| \\ &\leq \sum_i \sqrt{\langle M\rangle_{s_i}^{t_i} \langle N\rangle_{s_i}^{t_i}} \\ &\leq \left[\int_A d\langle M\rangle\right]^{1/2} \left[\int_A d\langle N\rangle\right]^{1/2}.\end{aligned}$$

La clase de unión ajena finita de intervalos de la forma $(s, t]$ es un álgebra que genera a $\mathscr{B}_{\mathbb{R}_+}$ y hemos mostrado que está contenida dentro de la clase

$$\mathscr{M} = \left\{A \in \mathscr{B}_{\mathbb{R}_+} : \int_A |d\langle M, N\rangle| \leq \left(\int_A d\langle M\rangle\right)^{1/2} \left(\int_A d\langle N\rangle\right)^{1/2}\right\},$$

que resulta ser una clase monótona. Por lo tanto $\mathscr{M} = \mathscr{B}_{\mathbb{R}_+}$.

Si H y K son dos procesos medibles de la forma

$$H = \sum \lambda_i \mathbf{1}_{A_i} \quad \text{y} \quad H = \sum \mu_i \mathbf{1}_{A_i}$$

donde λ_i y μ_i son variables aleatorias no-negativas y hemos representado a los procesos en términos de los mismos conjuntos A_i, entonces al aplicar la desigualdad del párrafo anterior y la de Cauchy obtenemos

$$\begin{aligned}\int H_r K_r \ |d\langle M,N\rangle_r| &= \sum_i \lambda_i \mu_i \int_{A_i} |d\langle M,N\rangle| \\ &\leq \sum_i \lambda_i \mu_i \left[\int_{A_i} |d\langle M\rangle|\right]^{1/2} \left[\int_{A_i} |d\langle M\rangle|\right]^{1/2} \\ &\leq \left[\sum_i \lambda_i^2 \int_{A_i} |d\langle M\rangle|\right]^{1/2} \left[\sum_i \mu_i^2 \int_{A_i} |d\langle N\rangle|\right]^{1/2} \\ &= \left[\int H_r \ |d\langle M\rangle_r|\right]^{1/2} \left[\int K_r \ |d\langle N\rangle_r|\right]^{1/2}.\end{aligned}$$

En particular la desigualdad anterior vale si los procesos H y K son funciones medibles simples en el espacio producto.

Para terminar la demostración, basta aproximar a los procesos medibles no-negativos H y K del enunciado por procesos simples. □

2.1. Construcción directa de la integral estocástica. Podemos entonces proceder con la construcción de la integral estocástica, primero en integrandos elementales.

Definición. Los **integrandos elementales** son los elementos de $L_2(M)$ que se escriben en la forma

$$H = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{(t_{i-1},t_i]}$$

donde $0 = t_0 < t_1 < \cdots < t_n$ y λ_i es $\mathscr{F}_{t_{i-1}}$-medible y acotada.

Denotaremos por $\mathscr{E}$ a la clase de procesos elementales.

Notemos que efectivamente un tal proceso pertenece a $L_2(M)$. En efecto, si C es una cota para H, entonces

$$\mathbb{E}\left(\int H_r^2 \, d\langle M\rangle_r\right) \leq C^2 \mathbb{E}(\langle M\rangle_\infty).$$

Por otra parte, puesto que $M^2 - \langle M\rangle$ es una martingala continua acotada en L_2 entonces es una martingala local y si T_n es una sucesión de tiempos de paro que la localiza, se tiene que

$$\mathbb{E}\big(M_{T_n}^2\big) = \mathbb{E}(\langle M\rangle_{T_n}).$$

Conforme $n \to \infty$, el lado derecho tiende a $\mathbb{E}(\langle M\rangle_\infty)$ por el teorema de convergencia monótona. Por otra parte, la desigualdad L_2 de Doob muestra que

$$\mathbb{E}\left(\sup_s M_s^2\right) \leq 4\mathbb{E}\big(M_\infty^2\big) < \infty,$$

por lo que podemos aplicar convergencia dominada para ver que

$$\lim_{n\to\infty} \mathbb{E}\big(M_{T_n}^2\big) = \mathbb{E}\big(M_\infty^2\big) .$$

Por lo tanto $\langle M\rangle_\infty \in L_1$,

$$\mathbb{E}(\langle M\rangle_\infty) = \mathbb{E}\big(M_\infty^2\big)$$

y

$$\mathbb{E}\left(\int H_r^2 \, d\langle M\rangle_r\right) \leq C^2\mathbb{E}(\langle M\rangle_\infty) < \infty,$$

por lo que $H \in L_2(M)$.

Proposición 3.4. *Si $M \in H_2$ entonces $\mathscr{E}$ es denso en $L_2(M)$.*

Demostración. Recordemos que $L_2(M)$ es un espacio de Hilbert, por lo que para probar la densidad de $\mathscr{E}$ basta probar que si $K \in L_2(M)$ es ortogonal a $\mathscr{E}$ entonces $H = 0$.

Sea

$$X_t = \int_0^t K_s \, d\langle M\rangle_s.$$

Si K es ortogonal a $\mathscr{E}$, probaremos que X es martingala. Puesto que X es continuo y tiene trayectorias de variación finita, esto nos dirá que X tiene trayectorias constantes, por lo que $K_s = 0$ para $\langle M\rangle$-casi toda $t \geq 0$ y para casi toda $\omega \in \Omega$. Así, $K = 0$ en $L_2(M)$.

Concluimos la prueba al mostrar que X es martingala. Si $A \in \mathscr{F}_s$ y $s < t$, definamos a $H_r(\omega) = \mathbf{1}_A(\omega)\,\mathbf{1}_{(s,t]}(r)$. Entonces $H \in \mathscr{E}$ y por lo tanto

$$0 = \mathbb{E}\left(\int_0^\infty H_r K_r \, d\langle M\rangle_r\right) = \mathbb{E}\left(\mathbf{1}_A \int_s^t K_r \, d\langle M\rangle_r\right) = \mathbb{E}(\mathbf{1}_A\,(X_t - X_s)) .$$

□

Definición. Sea $M \in H_2$. Se define la **integral estocástica elemental** respecto de M como la asignación de $\mathscr{E} \subset L_2(M)$ en H_2 que a $H = \sum_{i=1}^n \lambda_i \mathbf{1}_{(t_{i-1},t_i]}$ le asigna la martingala continua $H \cdot M$ dada por

$$(H \cdot M)_t = \sum_{i=1}^n \lambda_i \left(M_{t\wedge t_i} - M_{t\wedge t_{i-1}}\right) .$$

Notemos que en efecto $H \cdot M$ es una martingala acotada en L_2. Si C es una cota para H, entonces

$$(H \cdot M)_t \leq C M_t,$$

por lo que

$$\mathbb{E}\Big((H \cdot M)_t^2\Big) \leq C\mathbb{E}\big(M_\infty^2\big) .$$

Por otra parte, si $s < t$, digamos que $s \in (t_{i-1}, t_i]$ y $t \in (t_{j-1}, t_j]$. Entonces

$$\begin{aligned}
&\mathbb{E}\big((H \cdot M)_t - (H \cdot M)_s \mid \mathscr{F}_s\big) \\
&= \lambda_i \mathbb{E}\big(M_{t_{i+1}} - M_s \mid \mathscr{F}_s\big) + \sum_{l=i+2}^{n} \mathbb{E}\big(\lambda_l \left(M_{t\wedge t_l} - M_{t\wedge t_{l-1}}\right) \mid \mathscr{F}_s\big) \\
&= 0.
\end{aligned}$$

De hecho, podemos calcular exctamente la norma de $H \cdot M$ en H_2: notemos que, por ortogonalidad de los incrementos de las martingalas cuadrado integrables

$$\begin{aligned}
\|H \cdot M\|_{H_2}^2 &= \mathbb{E}\Big((H \cdot M)_\infty^2 \Big) \\
&= \mathbb{E}\left(\left[\sum_{i=1}^{n} \lambda_i \left(M_{t_i} - M_{t_{i-1}} \right) \right]^2 \right) \\
&= \mathbb{E}\left(\sum_{i=1}^{n} \lambda_i^2 \left(M_{t_i} - M_{t_{i-1}} \right)^2 \right).
\end{aligned}$$

Al utilizar aditividad de la esperenza, podemos calcular cada sumando al condicionar por $\mathscr{F}_{t_{i-1}}$, notando que

$$\begin{aligned}
\mathbb{E}\Big(\left(M_{t_i} - M_{t_{i-1}} \right)^2 \,\Big|\, \mathscr{F}_{t_{i-1}} \Big) &= \mathbb{E}\Big(M_{t_i}^2 - M_{t_{i-1}}^2 \,\Big|\, \mathscr{F}_{t_{i-1}} \Big) \\
&= \mathbb{E}\big(\langle M \rangle_{t_i} - \langle M \rangle_{t_{i-1}} \,\big|\, \mathscr{F}_{t_{i-1}} \big).
\end{aligned}$$

Se concluye que

$$\|H \cdot M\|_{H_2}^2 = \mathbb{E}\left(\sum_{i=1}^{n} \lambda_i^2 \left[\langle M \rangle_{t_i} - \langle M \rangle_{t_{i-1}} \right] \right) = \mathbb{E}\left(\int_0^\infty H_s \, d\langle M \rangle_s \right) = \|H\|_{L_2(M)}.$$

Vemos entonces que la asignación $H \mapsto H \cdot M$ preserva la norma.

De hecho, podemos probar aún más: que para cualqueir $N \in H_2$ se tiene que $\langle H \cdot M, N \rangle = H \cdot \langle M, N \rangle$. En efecto, probemos que si $s \leq t$ entonces

$$\mathbb{E}(N_t H \cdot M_t - N_s H \cdot M_s \mid \mathscr{F}_s) = \mathbb{E}(H \cdot \langle M, N \rangle_t - H \cdot \langle M, N \rangle_s \mid \mathscr{F}_s).$$

Al incluir a s y a t dentro de la partición que define a H, podemos suponer que $s = t_i < t_j = t$. Primero simplificamos al lado izquierdo al escribir

$$\begin{aligned}
&\mathbb{E}\big(N_{t_j} H \cdot M_{t_j} - N_{t_i} H \cdot M_{t_i} \,\big|\, \mathscr{F}_{t_i}\big) \\
&= \mathbb{E}\big(N_{t_j} H \cdot M_{t_j} - N_{t_i} H \cdot M_{t_j} + N_{t_i} H \cdot M_{t_j} - N_{t_i} H \cdot M_{t_i} \,\big|\, \mathscr{F}_{t_i}\big).
\end{aligned}$$

Puesto que $H \cdot M$ es martingala, vemos que entonces

$$\mathbb{E}\big(N_{t_j} H \cdot M_{t_j} - N_{t_i} H \cdot M_{t_i} \,\big|\, \mathscr{F}_{t_i}\big) = \mathbb{E}\big(N_{t_j} H \cdot M_{t_j} - N_{t_i} H \cdot M_{t_j} \,\big|\, \mathscr{F}_{t_i}\big).$$

Ahora sustituimos la definición de la integral estocástica elemental para obtener

$$\mathbb{E}\big(N_{t_j} H \cdot M_{t_j} - N_{t_i} H \cdot M_{t_j} \,\big|\, \mathscr{F}_{t_i} \big) = \mathbb{E}\left(\sum_{l=1}^{j} \left(N_{t_j} - N_{t_i}\right) \lambda l \left(M_{t_l} - M_{t_{l-1}}\right) \,\middle|\, \mathscr{F}_{t_i} \right).$$

Al utilizar el carácter martingala de M se obtiene:

$$\mathbb{E}\big(\left(N_{t_j} - N_{t_i}\right) \lambda_l \left(M_{t_l} - M_{t_{l-1}}\right) \,\big|\, \mathscr{F}_{t_i} \big) = \begin{cases} 0 & l \leq i \\ \mathbb{E}\big(N_{t_j} \lambda_l \left(M_{t_l} - M_{t_{l-1}}\right) \,\big|\, \mathscr{F}_{t_i} \big) & i < l \leq j \end{cases}.$$

Así:

$$\mathbb{E}\big(N_{t_j} H \cdot M_{t_j} - N_{t_i} H \cdot M_{t_i} \,\big|\, \mathscr{F}_{t_i} \big) = \sum_{l=i+1}^{j} \mathbb{E}\big(N_{t_j} \lambda_l \left(M_{t_l} - M_{t_{l-1}}\right) \,\big|\, \mathscr{F}_{t_i} \big).$$

Al separar el incremento y condicionar por $\mathscr{F}_{t_l}$ y por $\mathscr{F}_{t_{l-1}}$ obtenemos

$$\mathbb{E}\big(N_{t_j} \lambda_l \left(M_{t_l} - M_{t_{l-1}}\right) \,\big|\, \mathscr{F}_{t_i} \big) = \mathbb{E}(N_{t_l} \lambda_l M_{t_l} \mid \mathscr{F}_{t_i}) - \mathbb{E}\big(N_{t_{l-1}} \lambda_l M_{t_{l-1}} \,\big|\, \mathscr{F}_{t_i} \big).$$

Finalmente, al utilizar la propiedad característica de la covariación y al condicionar por $\mathscr{F}_{t_{l-1}}$ se obtiene

$$\mathbb{E}(N_{t_l} \lambda_l M_{t_l} \mid \mathscr{F}_{t_i}) - \mathbb{E}\big(N_{t_{l-1}} \lambda_l M_{t_{l-1}} \,\big|\, \mathscr{F}_{t_i} \big) = \mathbb{E}(\lambda_l \langle M, N \rangle_{t_l} \mid \mathscr{F}_{t_i}).$$

Se concluye que

$$\begin{aligned} \mathbb{E}\big(N_{t_j} H \cdot M_{t_j} - N_{t_i} H \cdot M_{t_i} \,\big|\, \mathscr{F}_{t_i} \big) &= \sum_{l=i+1}^{j} \mathbb{E}(\lambda_l \langle M, N \rangle_{t_l} \mid \mathscr{F}_{t_i}) \\ &= \mathbb{E}\big(H \cdot \langle M, N \rangle_{t_j} - H \cdot \langle M, N \rangle_{t_i} \,\big|\, \mathscr{F}_{t_i} \big). \end{aligned}$$

Teorema 3.3. *La asignación $H \mapsto H \cdot M$ se extiende a una isometría de $L_2(M)$ en el subespacio de los elementos de H_2 que comienzan en cero. Además $H \cdot M$ es único elemento de H_2 nulo en cero tal que para toda $N \in H_2$*

$$\langle H \cdot M, N \rangle = H \cdot \langle M, N \rangle.$$

En particular $\langle H \cdot M \rangle = H^2 \cdot \langle M \rangle$. Finalmente, si T es un tiempo de paro el proceso $\mathbf{1}_{[0,T]}$ definido por $(t, \omega) \mapsto \mathbf{1}_{T(\omega) \leq t}$ pertenece a $L_2(M)$ y

$$\big(\mathbf{1}_{[0,T]} H\big) \cdot M = (H \cdot M)^T = H \cdot M^T.$$

Demostración. Notemos que $H \mapsto H \cdot M$ es una aplicación lineal de $\mathscr{E}$ en H_2. (Más precisamente, su imagen consta de martingalas que comienzan en cero.)

Si $H \in L_2(M)$, sea H_n una sucesión de elementos de $\mathscr{E}$ que convergen a H en $L_2(M)$. Entonces

$$\| H_n \cdot M - H_m \cdot M \|_{H_2} = \| \left(H_n - H_m\right) \cdot M \|_{H_2} = \| H_n - H_m \|_{L_2(M)}$$

por lo que la sucesión $(H_n \cdot M)$ es de Cauchy en H_2. Al ser H_2 un espacio de Hilbert, existe un elemento $H \cdot M$ de H_2 tal que $H_n \cdot M \to H \cdot M$. Por otra parte, si H_n^1 y H_n^2 son dos sucesiones que convergen a H en $L_2(M)$ entonces

$$\|H_n^1 \cdot M - H_n^2 \cdot M\|_{H_2} = \|H_n^1 - H_n^2\|_{L_2(M)} \to 0.$$

Concluimos que $H \cdot M$ no depende de la sucesión de elementos que aproximen a H y entonces $H \cdot M$ es una extensión de la integral estocástica elemental a $L_2(M)$ heredando así la linealidad y la propiedad de ser isometría.

Si ahora $N, M \in H_2$ y $H \in L_2(M)$, sea H_n una sucesión de procesos en $\mathscr{E}$ que converge a H en $L_2(M)$. Puesto que $H_n \cdot M \to H \cdot M$ en H_2, vemos que $(H_n \cdot M)_\infty^2 N_\infty^2$ converge a $(H \cdot M)_\infty^2 N_\infty^2$ en L_1. Por otra parte, de la desigualdad de Kunita-Watanabe se obtiene

$$\begin{aligned}\mathbb{E}(|[(H_n - H) \cdot \langle M, N\rangle]_t|) &\leq \mathbb{E}\Big([|H_n - H| \cdot \langle M\rangle]_\infty^{1/2} \langle N\rangle_\infty^{1/2}\Big)\\ &\leq \|\,(H_n - H) \cdot M\|_{H_2}\|N\|_{H_2},\end{aligned}$$

por lo que $(H_n \cdot \langle M, N\rangle)_t$ converge a $(H \cdot \langle M, N\rangle)_t$ en L_1. Al pasar al límite, se deduce entonces que $(H \cdot M)^2 N^2 - H \cdot \langle M, N\rangle$ es una martingala, por lo que $\langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle$.

Sea $\tilde{M} \in H_2$ nula en cero y tal que

$$\langle \tilde{M}, N\rangle = H \cdot \langle M, N\rangle$$

para toda $N \in H_2$. Entonces

$$\langle \tilde{M}\rangle = H^2 \cdot \langle M\rangle = \langle H \cdot M\rangle.$$

Sin embargo, también de deduce que

$$\langle \tilde{M}, H \cdot M\rangle = H \cdot \langle M, H \cdot M\rangle = H^2\langle M\rangle,$$

por lo que se deduce que

$$\langle \tilde{M} - H \cdot M\rangle = 0.$$

Así, M y $H \cdot M$ son indistinguibles.

Finalmente,si T es un tiempo de paro entonces $\mathbf{1}_{[0,T]} \in L_2(M)$. Por una parte, vemos que

$$\langle \mathbf{1}_{[0,T]} H \cdot M, N\rangle = \mathbf{1}_{[0,T]} H \cdot \langle M, N\rangle = H \cdot \langle M, N\rangle^T = H \cdot \langle M^T, N\rangle.$$

Se concluye que $\mathbf{1}_{[0,T]} H \cdot M = H \cdot M^T$. Además, vemos que

$$\langle (H \cdot M)^T, N\rangle = \langle H \cdot M, N\rangle^T = H \cdot \langle M, N\rangle^T = H \cdot \langle M^T, N\rangle,$$

por lo que también $\mathbf{1}_{[0,T]} H \cdot M = (H \cdot M)^T$. □

Las siguientes propiedades son útiles para probar propiedades de la integral estocástica y ampliar a los posibles integrandos e integradores.

Proposición 3.5. *Si $M \in H_2$, $H \in L_2(M)$ y $K \in L_2(H \cdot M)$ entonces $HK \in L_2(M)$ y $K \cdot (H \cdot M) = KH \cdot M$.*

Demostración. Notemos que por el Teorema 3.3

$$\int_{[0,\infty)} K^2 H^2 \, d\langle M\rangle = \int_{[0,\infty)} K^2 \, d\langle H \cdot M\rangle.$$

Se concluye entonces que $KH \in L_2(M)$.

Por otra parte, notemos que si $N \in H_2$ entonces

$$\langle K \cdot (H \cdot M), N\rangle = K \cdot \langle H \cdot M, N\rangle = K \cdot (H \cdot \langle M, N\rangle) = KH \cdot \langle M, N\rangle.$$

Por el Teorema 3.3 concluimos que $K \cdot (H \cdot M) = KH \cdot M$. □

2.2. Construcción de la integral estocástica mediante el teorema de representación de Riesz.

3. La integral estocástica respecto de semimartingalas

En esta sección extenderemos en dos etapas el resultado de existencia de la integral estocástica al caso de las semimartingalas (pasando primero por martingalas locales) mediante el método de localización.

Sea M una martingala local continua y

$$T_n = \inf\left\{t \geq 0 : |M_t| \geq n \text{ ó } \langle M\rangle_t \geq n\right\}.$$

Notemos que $M^{T_n} \in H_2$.

Comencemos con la definición del espacio de integrandos.

Definición. El espacio $L_2^{\mathrm{loc}}(M)$ se define como la clase de procesos estocásticos progresivamente medibles K para los cuales existe una sucesión de tiempos de paro (S_n) tal que $S_n \leq S_{n+1}$, $S_n \to \infty$ casi seguramente y

$$\mathbb{E}\left(\int_0^{S_n} K_s^2 \, d\langle M\rangle_s\right) < \infty.$$

Si K es continuo y adaptado, entonces $K \in L_2^{\mathrm{loc}}(M)$. Esto sucede pues si

$$S_n = \inf\left\{t \geq 0 : |K_s| \geq n \text{ ó } \langle M\rangle_s > n\right\}$$

entonces por continuidad tanto K como $\langle M\rangle$ son acotados en $[0, n \wedge S_n]$ y entonces

$$\mathbb{E}\left(\int_0^{n\wedge S_n} K_s^2 \, d\langle M\rangle_s\right) \leq n^4.$$

De hecho, este ejemplo es un caso particular del siguiente.

Definición. Un proceso progresivamente medible K es **localmente acotado** si existe una sucesión de tiempos de paro (S_n) tal que $S_n \leq S_{n+1}$, $S_n \to \infty$ casi seguramente y K^{S_n} es acotado.

Ejercicio 3.2. Pruebe que si M es una martingala local continua y K es localmente acotado entonces $K \in L_2^{\mathrm{loc}}(M)$. Dé un ejemplo de proceso localmente acotado que no sea continuo.

Teorema 3.4. *Sea M una martingala local continua y $H \in L_2^{loc}(M)$. Entonces existe una única martingala local continua que se anula en cero, denotado $H \cdot M$ tal que para cualquier martingala local continua N*

$$\langle H \cdot M, N \rangle = H \cdot \langle M, N \rangle .$$

Al proceso $H \cdot M$ se le da el nombre de **integral estocástica** de H respecto de M.

Demostración. Respecto a la existencia, considere una sucesión (S_n) de tiempos de paro como en la definición de $L_2^{\mathrm{loc}}(M)$. Puesto que $M^{S_n \wedge T_n} \in H_2$ y $H^{S_n \wedge T_n} \in L_2(M)$, podemos considerar a $I_n = H^{S_n \wedge T_n} \cdot M^{S_n \wedge T_n}$. Por la propiedad de localización, vemos que $I_{n+1}^{S_n \wedge T_n} = I_n$ y como $S_n \wedge T_n \to \infty$ casi seguramente, podemos definir a $H \cdot M$ mediante

$$(H \cdot M)_t = \lim_{n \to \infty} I_n(t) .$$

Por construcción, $(H \cdot M)^{S_n} = I_n$, por lo que la propiedad característica de la integral estocástica nos dice que si N es una martingala local continua y

$$R_n = \inf \{ t \geq 0 : |N_t| \geq n \}$$

entonces

$$\langle H \cdot M, N \rangle^{S_n \wedge T_n \wedge R_n} = \langle I_n^{S_n \wedge T_n}, N^{R_n} \rangle = H^{S_n \wedge T_n} \cdot \langle M^{S_n \wedge T_n}, N^{R_n} \rangle \to H \cdot \langle M, N \rangle .$$

(Note que sobre el conjunto $t < S_n \wedge T_n \wedge R_n$, la integral $\int_0^t H_s \, d\langle M, N \rangle_s$ está bien definida. Como la unión sobre n de dichos conjuntos tiene probabilidad 1, se ha mostrado que la integral de Riemann-Stieltjes está bien definida casi seguramente.)

La unicidad se sigue del siguiente argumento: si I es una martingala local continua nula en cero tal que para cualquier otra martingala local continua N se satisface $\langle I, N \rangle = H \cdot \langle M, N \rangle$ entonces $\langle I \rangle = H^2 \cdot \langle M \rangle = \langle I, H \cdot M \rangle$. De aquí se concluye que $\langle I - H \cdot M \rangle = \langle I \rangle + \langle H \cdot M \rangle - 2 \langle I, H \cdot M \rangle = 0$. Puesto que $I - H \cdot M$ es nula en cero, la Proposición 3.3 nos dice que I y $H \cdot M$ son indistinguibles. □

La integral estocástica satisface la propiedad de detención y de aditividad que en el caso de martingalas acotadas en L_2. La prueba se deja como **ejercicio**.

Pasemos ahora al caso de las semimartingalas. Sea X una martingala local continua con descomposición canónica $X = X_0 + M + A$ donde M es una martingala local continua (nula en cero) y A es un proceso continuo de variación acotada en compactos. El espacio adecuado de integrandos lo conformaran los procesos predecibles localmente acotados. Si K es un tal proceso, se define la **integral estocástica** de K respecto de X, denotada por $K \cdot X$, como el proceso estocástico dado por

$$(K \cdot X)_t = K \cdot M + K \cdot A .$$

Notemos que $K \cdot X$ es una nueva semimartingala. El siguiente resultado resume las propiedades más importantes de la integral estocástica.

Teorema 3.5. *Sean $X = X_0 + M + A$ una semimartingala continua, y H, H_n y K procesos progresivamente medibles localmente acotados. Entonces*

(1) $K \cdot (H \cdot X) = KH \cdot X$,
(2) *si T es un tiempo de paro entonces $\mathbf{1}_{[0,T]} H \cdot X = (H \cdot X)^T = H \cdot X^T$,*
(3) *si $H \in \mathscr{E}$ tiene descomposición $H = \sum \lambda_i \mathbf{1}_{[t_{i-1},t_i]}$ entonces*

$$(H \cdot X)_t = \sum_i \lambda_i \left(X_{t_i} - X_{t_{i-1}} \right)$$

(4) *si $H_n \to H$ uniformemente en compactos en probabilidad y $|H_n| \leq K$ entonces $H_n \cdot X \to H \cdot X$ uniformemente en compactos en probabilidad,*
(5) *y si H es continuo por la derecha y Δ_n es una sucesión de particiones de $[0, t]$ cuya norma tiende a cero entonces*

$$\int_0^t H_s \, dX_s = \lim_{n \to \infty} \sum_{\Delta_n} H_{t_{i-1}} \left(X_{t_i} - X_{t_{i-1}} \right).$$

en probabilidad.

Demostración. Las primeras 3 propiedades son válidas para ambas integrales por lo que lo son para la integral estocástica respecto de semimartingalas.

Supongamos que $H_n \to H$ uniformemente en compactos en probabilidad y que $|H_n| \leq K$. Sea

$$S_n = \inf \left\{ s \geq 0 : K_s \geq n/2 \text{ ó } \langle M \rangle_s \geq n \text{ ó } \ |M_s| \geq n \right\} \wedge t.$$

Notemos que $\lim_{n \to \infty} S_n = t$. Fijemos $\delta > 0$ y sea n tal que

$$\mathbb{P}(S_n < t) \leq \delta$$

y notemos que $|H_m(s) - H(s)| \leq n$ para toda m si $s \leq S_n$. Puesto que $M^{S_n} \in H_2$ y $(H - H_m)^{S_n} \in L_2\big(M^{S_n}\big)$ vemos que

$$\mathbb{E}\left(\sup_{s \leq S_n} \left| (H \cdot M)_s - (H_m \cdot M)_s \right|^{2} \right) \leq 4 \mathbb{E}\left(\int_0^{S_n} |H - H_m|_s^2 \, d\langle M \rangle_s \right) \to_{m \to \infty} 0$$

por el teorema de convergencia acotada pues

$$. \int_0^{S_n} |H - H_m|_s^2 \, d\langle M \rangle_s \leq n^3 t.$$

Puesto que

$$\begin{aligned} &\mathbb{P}\left(\sup_{s \leq t} \left| (H \cdot M)_s - (H_m \cdot M)_s \right| \right) \\ &\leq \mathbb{P}(S_n^m < t) + \frac{1}{\varepsilon^2} \mathbb{E}\left(\sup_{s \leq S_n^m} \left| (H \cdot M)_s - (H_m \cdot M)_s \right|^{2} \right), \end{aligned}$$

vemos que

$$\limsup_{m\to\infty} \mathbb{P}\left(\sup_{s\leq t} \left|(H\cdot M)_s - (H_m\cdot M)_s\right|\right) \leq \delta$$

para toda $\delta > 0$, por lo que de hecho $H_m \cdot M \to H \cdot M$ uniformemente en compactos en probabilidad.

Si H es continuo por la izquierda, podemos aplicar el resultado del párrafo anterior a la sucesión H_n dada por

$$H_n(t) = \sum_{\Delta_n} H(t_{i-1})\, \mathbf{1}_{[t_{i-1},t_i]}(t)$$

que está acotada por H. $\square$

4. La fórmula de Itô

En esta sección probaremos una versión estocástica de la regla de la cadena para la integral de Riemann-Stieltjes. De hecho, ya hemos probado una versión para el movimiento browniano, misma que generalizaremos a martingalas continuas. La demostración que daremos no será una generalización de la que ya vimos para el movimiento browniano sino que, como en el caso de la integral de Riemann-Stieltjes, nos basaremos en (una versión estocástica de) la fórmula de integración por partes.

Teorema 3.6 (Fórmula de integración por partes). *Sean X y Y semimartingalas continuas. Entonces*

$$XY = X_0Y_0 + X\cdot Y + Y\cdot X + \langle X, Y\rangle.$$

En particular, $X^2 = 2X\cdot X + \langle X\rangle$.

Si M y N son dos martingalas locales continuas, ya sabíamos que $MN - \langle M, N\rangle$ es una martingala local continua. Lo que nos da la fórmula de integración por partes es una representación explícita de dicha martingala en términos de integrales estocásticas.

Demostración. Se hará por aproximación. Sea Δ_n una sucesión de particiones de $[0,t]$ cuyo paso tiende a cero. Puesto que

$$\begin{aligned} X_tY_t - X_0Y_0 &= \sum_{\Delta_n} \left(X_{t_i}Y_{t_i} - X_{t_{i-1}}Y_{t_{i-1}}\right) \\ &= \sum_{\Delta_n} Y_{t_{i-1}}\left(X_{t_i} - X_{t_{i-1}}\right) + X_{t_{i-1}}\left(Y_{t_i} - Y_{t_{i-1}}\right) \\ &\quad + \left(X_{t_i} - X_{t_{i-1}}\right)\left(Y_{t_i} - Y_{t_{i-1}}\right), \end{aligned}$$

podemos tomar el límite conforme $n \to \infty$ para concluir la fórmula de integración por partes, gracias a los teoremas de aproximación de la integral estocástica y de la covariación. $\square$

Si $X = \left(X^1, \ldots, X^d\right)$ es un proceso estocástico con valores en $\mathbb{R}^d$ tal que cada componente es una semimartingala continua, decimos que X es una **semimartingala vectorial**. Si $F : \mathbb{R}^d \to \mathbb{R}$ es dos veces diferenciable y $e^i \in \mathbb{R}^e$ denota al i-ésimo vector de la base canónica que tiene todas las entradas iguales a cero salvo la i-ésima igual a 1, denotaremos por $D_i F$ a la derivada de F en la dirección e^i. La notación $D_{i,j}F$ se utilizará para $D_j(D_i F)$, misma que se abreviará como D_i^2 cuando $i = j$. Cuando $d = 1$, se utiliza la notación D y D^2.

Teorema 3.7 (Fórmula de Itô). *Sea $X = \left(X^1, \ldots, X^d\right)$ una semimartingala vectorial y $F : \mathbb{R}^d \to \mathbb{R}$ de clase C_2. Entonces el proceso $F(X) = (F(X_t))_{t\geq 0}$ es una semimartingala real con descomposición*

$$F(X) = F(X_0) + \sum_{i=1}^{d} D_i F(X_s) \cdot X^i + \frac{1}{2} \sum_{i,j=1}^{d} D_{i,j} F(X) \cdot \langle X^i, X^j \rangle.$$

Esta descomposición de $F(X)$ se conoce con el nombre de fórmula de Itô y usualmente se escribe de la siguiente manera:

$$F(X_t) = F(X_0) + \sum_{i=1}^{d} \int_0^t \frac{\partial F}{\partial x_i}(X_s)\, dX_s^i + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \frac{\partial F}{\partial x_i \partial x_j}(X_s)\, d\langle X^i, X^j \rangle_s.$$

Demostración. Haremos primero el caso $d = 1$. Primero notamos que si la fórmula de Itó se satisface para F y $\alpha, \beta \in \mathbb{R}$ entonces se satisface para $\alpha F + \beta$ y para la función $G : \mathbb{R} \to \mathbb{R}$ dada por

$$G(x) = xF(x)$$

Supongamos que

$$F(X) = F(X_0) + DF(X) \cdot X + \frac{1}{2} D^2 F(X) \cdot \langle X \rangle.$$

La fórmula de integración por partes nos dice que

$$XF(X) = X_0 F(X_0) + X \cdot F(X) + F(X) \cdot X + \langle X, F(X) \rangle.$$

Al utilizar la descomposición que hemos supuesto para $F(X)$, vemos que

$$\begin{aligned} G(X) &= X_0 F(X_0) + XDF(X) \cdot X + \frac{1}{2} X D^2 F(X) \cdot \langle X \rangle + F(X) \cdot X + \langle X, F(X) \rangle \\ &= G(X_0) + DG(X) \cdot X + \frac{1}{2} D^2 G(X) \cdot \langle X \rangle. \end{aligned}$$

Por lo tanto, la fórmula de Itô se satisface para G y entonces la fórmula de Itô se satisface para polinomios. Sea $I \subset \mathbb{R}$ un intervalo compacto y $T = \inf \{t \geq 0 : X_t \notin I\}$. Entonces T_I es un tiempo de paro y X^I es una semimartingala continua que permanece en I. Puesto que $F \in C_2$, existe una sucesión de polinomios $p_k, k \geq 1$ tal que $D^2 p_k$, Dp_k y p_k convergen uniformemente en I conforme $k \to \infty$ hacia $D^2 F, DF$ y F. Al utilizar el teorema de convergencia acotada para la integral de

Lebesgue-Stieltjes y el teorema de continuidad para la integral estocástica, vemos que la fórmula de Itô se satisface para F. El caso d-dimensional sólo difiere en necesitar una notación más complicada. $\square$

CAPÍTULO 4

Aplicaciones a la integral estocástica

En este capítulo estudiaremos las aplicaciones a la teoría de integración estocástica desarrollada en el capítulo anterior.

1. La exponencial estocástica

Comencemos con la construcción de la exponencial estocástica de una martingala local continua M.

TEOREMA 4.1. *Existe un único proceso continuo y adaptado $\mathscr{E}(M)$ tal que*

$$\mathscr{E}(M)_t = 1 + \int_0^t \mathscr{E}(M)_s \, dM_s.$$

Se tiene la fórmula explícita

$$\mathscr{E}(M)_t = e^{M_t - \frac{1}{2}\langle M \rangle_t}.$$

DEMOSTRACIÓN. Sea $\mathscr{E}(M)_t = e^{M_t - \frac{1}{2}\langle M \rangle_t}$. Al aplicar la fórmula de Itô con la función $f(x_1, x_2) = e^{x - y/2}$ y con la semimartingala vectorial $X = (M, \langle M \rangle)$, vemos que

$$\mathscr{E}(M)_t = 1 + \int_0^t \mathscr{E}(M)_s \, dM_s - \int_0^t \frac{1}{2} \mathscr{E}(M)_s \, d\langle M \rangle_s + \frac{1}{2} \int_0^t \mathscr{E}(M)_s \, d\langle M \rangle_s,$$

por lo que $\mathscr{E}(M)$ satisface la ecuación diferencial estocástica anunciada.

Por otra parte, notemos que $\mathscr{E}(M) > 0$, por lo que podemos aplicar la fórmula de Itô y concluir que

$$\mathscr{E}(M)_t^{-1} = 1 - \int_0^t \mathscr{E}(M)_s^{-1} \, dM_s + \int_0^t \mathscr{E}(M)_s^{-1} \, d\langle M \rangle_s.$$

Supongamos que X es continuo, adaptado y satisface la ecuación diferencial estocástica

$$X_t = 1 + \int_0^t X_s \, dM_s.$$

Entonces la fórmula de integración por partes nos permite deducir que

$$\begin{aligned} X_t\mathscr{E}(M)_t^{-1} &= 1 + \int_0^t X_s\mathscr{E}(M)_s^{-1}\,(d\langle M\rangle_s - M_s) \\ &\quad + \int_0^t X_s\mathscr{E}(M)_t^{-1}\,dM_s - \int_0^t X_s\mathscr{E}(M)_t^{-1}\,d\langle M\rangle_s \\ &= 1. \end{aligned}$$

Por lo tanto, concluimos que $X = \mathscr{E}(M)$. □

2. El teorema de caracterización de Lévy

En esta sección haremos una primera aplicación de la fórmula de Itô a los procesos estocásticos. Recordemos que un proceso de Lévy es un proceso con trayectorias càdlàg que comienza en cero y que tiene incrementos independientes y estacionarios. Como en el caso browniano, se prueba que si el proceso de Lévy está definido en un espacio de probabilidad completo entonces su filtración completada satisface las hipótesis habituales. El problema que nos plantearemos es el de caracterizar a todos los procesos de Lévy con trayectorias continuas. Sea X un tal proceso de Lévy. Ahora veremos que entonces $X_t \in L_p$ para toda $p \geq 1$ y que por lo tanto $\mathbb{E}(X_t) = t\mu$ donde $\mu = \mathbb{E}(X_1)$, que $\mathrm{Var}(M_t) = t\sigma^2$ donde $\sigma^2 \geq 0$ y que por lo tanto el proceso M dado por $M_t = X_t - t\mu$ es una martingala continua con variación cuadrática $\langle M\rangle_t = \sigma^2 t$. El teorema de caracterización de Lévy, que enunciaremos a continuación nos permite entonces verificar que si $\sigma > 0$ entonces M/σ es un movimiento browniano. Así, todos los procesos de Lévy con trayectorias continuas tienen la forma $\sigma B_t + \mu t$ donde B es un movimiento browniano. Probemos ahora que $X_t \in L_p$ para toda $p \geq 1$. En efecto, sean

$$T_0 = t \quad \text{y} \quad T_{n+1} = \inf\{t \geq T_n : |X_t - X_{T_n}| \geq 1\}.$$

Entonces $T_1 > 0$ casi seguramente pues X tiene trayectorias continuas. Por otra parte, la propiedad de Markov fuerte de X nos dice que las variables $T_1 - T_0, T_2 - T_1, \ldots$ son independientes e idénticamente distribuidas. Así, si definimos

$$\alpha = \mathbb{E}\big(e^{-T_1}\big),$$

se tiene que $\alpha \in [0,1)$ (pues $\alpha = 1$ diría que $T_1 = 0$ casi seguramente) y

$$\mathbb{E}\big(e^{-T_n}\big) = \alpha^n.$$

Por otra parte, notemos que por definición $\left|X^{T_n}\right| \leq n$ y por lo tanto

$$\mathbb{P}(|X_t| \geq n) \leq \mathbb{P}(T_n \leq t) \leq e^t\alpha^n.$$

Finalmente, puesto que para $p \geq 1$,

$$\mathbb{E}(|X_t|^{\,p}) = \int_0^\infty px^{p-1}\mathbb{P}(|X_t| \geq x) \leq \sum_{n\geq 0} p\,(n+1)^{p-1}\,e^t\alpha^n < \infty.$$

Veamos ahora que $\mathbb{E}(X_t) = t\mathbb{E}(X_1)$ y que $\mathrm{Var}(X_t) = t\mathrm{Var}(X_1)$. El argumento es el mismo en ambos casos y se reduce a la linealidad de soluciones continuas por la derecha a la ecuación funcional de Cauchy. Esto es, notemos que de hecho hemos probado que $\sup_{s\leq t}\mathbb{E}(|X_s|^p) < \infty$ para toda $p \geq 1$. Esto nos dice que las colecciones de variables aleatorias $\{X_s : s \leq t+1\}$ y $\left\{X_s^2 : s \leq t+1\right\}$ son uniformemente integrables. Puesto que X es continuo por la derecha, entonces las funciones $s \mapsto \mathbb{E}(X_s)$ y $s \mapsto \mathrm{Var}(X_s)$ son continuas por la derecha. Por otra parte, por linealidad de la esperenza y puesto que los incrementos son estacionarios, vemos que

$$\mathbb{E}(X_{t+s}) = \mathbb{E}(X_t) + \mathbb{E}(X_s)\,.$$

Además, puesto que los incrementos son estacionarios e independientes, observamos que

$$\mathrm{Var}(X_{t+s}) = \mathrm{Var}(X_t) + \mathrm{Var}(X_s)\,.$$

Así, se tiene que $\mathbb{E}(X_t) = t\mathbb{E}(X_1)$ y $\mathrm{Var}(X_t) = t\mathrm{Var}(X_1)$.

Teorema 4.2 (Teorema de caracterización de Lévy). *Sea M una martingala local continua con variación cuadrática $\langle M\rangle_t = t$. Entonces M es un $(\mathscr{F}_t)$-movimiento browniano.*

Demostración. Aplicaremos la versión compleja de la martingala exponencial. Esto es, para $u \in \mathbb{R}$, consideremos a

$$\mathscr{E}(iuM_t) = e^{iuM_t + u^2 t/2},$$

que es una martingala local compleja. Es decir, su parte real y su parte imaginaria son martingalas locales como se puede verificar facilmente. Por tener trayectorias acotadas en compactos, vemos que $\mathscr{E}(iuM)$ es una martingala compleja (y no sólo local). Por lo tanto, se sigue que para $s < t$:

$$\mathbb{E}\left(e^{iuM_t + u^2 t/2} \,\middle|\, \mathscr{F}_s\right) = e^{iuM_s + u^2 s/2},$$

por lo cual para todo $A \in \mathscr{F}_s$

$$\mathbb{E}\left(\mathbf{1}_A e^{iu(M_t - M_s)}\right) = \mathbb{E}(\mathbf{1}_A)\, e^{-u^2(t-s)/2}.$$

Se sigue que si $\mathbb{P}(A) > 0$, entonces bajo la medida $\mathbb{P}(\,\cdot \mid A)$, $M_t - M_s$ tiene la misma función característica que una variable gaussiana centrada de varianza $t - s$ y por lo tanto la misma distribución. Así, para todo $C \in \mathscr{B}_{\mathbb{R}}$:

$$\mathbb{P}(A, M_t - M_s \in C) = \mathbb{P}(A)\,\mathbb{P}(B_{t-s} \in C)\,.$$

Notemos que la fórmula anterior sigue siendo válida si $\mathbb{P}(A) = 0$. Se concluye que $M_t - M_s$ es independiente de $\mathscr{F}_s$ y que M es un $(\mathscr{F}_t)$-movimiento browniano. □

3. Martingalas locales continuas como cambios de tiempo del movimiento browniano

El objetivo de esta sección es probar que toda martingala local continua un movimiento browniano cambiado de tiempo.

Para esto, comenzaremos primero con nociones generales sobre cambios de tiempo.

Definición. Sea $C = (C_t, t \geq 0)$ un proceso estocástico con trayectorias càd y no decrecientes. Decimos que C es un **cambio de tiempo** si C_t es un $(\mathscr{F}_s)$-tiempo de paro para toda $t \geq 0$.

Podemos entonces definir a la filtración cambiada de tiempo $(\hat{\mathscr{F}}_t, t \geq 0)$ donde $\hat{\mathscr{F}}_t = \mathscr{F}_{C_t}$. Notemos que si X es un proceso $(\mathscr{F}_s)$-progresivo entonces el proceso cambiado de tiempo $\hat{X}$ dado por $\hat{X}_t = X_{C_t}$ es adaptado a $(\hat{\mathscr{F}}_t)$.

Proposición 4.1. *Si $(\mathscr{F}_s)$ satisface las condiciones habituales entonces $(\hat{\mathscr{F}}_t)$ también las satisface.*

Demostración. Si $\mathscr{N}$ es la clase de los conjuntos $\mathbb{P}$-nulos entonces $\mathscr{N} \subset \mathscr{F}_s$ para toda $s \geq 0$. Si $A \in \mathscr{N}$ entonces

$$A \cap \{C_t \leq s\} \subset A,$$

por lo que $A \cap \{C_t \leq s\} \in \mathscr{F}_s$ y por lo tanto $\mathscr{N} \subset \hat{\mathscr{F}}_t$.

Notemos que

$$\bigcap_{\varepsilon > 0} \hat{\mathscr{F}}_{t+\varepsilon} = \bigcap_{\varepsilon > 0} \{A \in \mathscr{F} : A \cap \{C_{t+\varepsilon} \leq s\} \in \mathscr{F}_s \text{ para toda } s \geq 0\}.$$

Puesto que C es continuo por la derecha, si $A \in \hat{\mathscr{F}}_{t+\varepsilon}$ para toda $\varepsilon > 0$ entonces

$$A \cap \{C_t < s\} = \bigcup_{n=1}^{\infty} A \cap \left\{C_{t+1/n} < s\right\}.$$

Por continuidad por la derecha de $(\mathscr{F}_s)$, se sigue que $A \cap \{C_t \leq s\} \in \mathscr{F}_s$ por lo que $A \in \hat{\mathscr{F}}_t$. Se concluye que $(\hat{\mathscr{F}}_t)$ es continua por la derecha. □

Ahora veremos cómo se transforman las martingalas locales continuas y las integrales estocásticas ante cambios de tiempo. Para esto, notemos que una martingala local continua M no necesariamente se convierte en otra ante cambios de tiempo. Por ejemplo, si $C_t = \inf\{s \geq 0 : M_s \geq t\}$ (donde M es una martingala local continua que oscila como el movimiento browniano, para que $C_t < \infty$ casi seguramente) vemos que $M_{C_t} = t$. Este último proceso claramente no es una martingala. Es más, puesto que los cambios de tiempo en general son discontinuos, vemos que el proceso $M \circ C$ podría serlo también. La siguiente condición técnica, que evita estas dos situaciones, resulta ser suficiente para que una martingala local continua se convierta en otra ante cambios de tiempo.

Definición. Sea X un proceso estocástico y C un cambio de tiempo. Decimos que X es **C-continuo** si X es constante en el intervalo $[C_{t-}, C_t]$ para toda $t > 0$.

Teorema 4.3. *Sea C un cambio de tiempo tal que $C_t < \infty$ para toda $t \geq 0$ casi seguramente y sea M una martingala local continua C-continua. Entonces $\hat{M}$ es una $(\hat{\mathscr{F}}_t)$-martingala local continua. Si N es otra martingala local continua C-continua entonces la covariación entre $\hat{M}$ y $\hat{N}$ es $\widehat{\langle M, N\rangle}$. En particular, $\langle \hat{M}\rangle = \widehat{\langle M\rangle}$. Si H es un elemento de $L_2^{loc}(M)$ entonces $\hat{H} \in L_2^{loc}(\hat{M})$ y*

$$\hat{H} \cdot \hat{M} = \widehat{H \cdot M}.$$

Demostración. Puesto que M es C-continua se sigue que $\hat{M}$ es un proceso estocástico continuo y adaptado. Si T es un $(\mathscr{F}_t)$-tiempo de paro que acota a M, definamos a

$$\hat{T} = \inf\left\{t \geq 0 : C_t > T\right\}.$$

Puesto que M^T es una martingala local continua y acotada, se sigue que es una martingala continua y acotada por lo que

$$\mathbb{E}\Big(M_{C_t \wedge T} \,\Big|\, \hat{\mathscr{F}}_s \Big) = M_{C_s \wedge T}.$$

Puesto que M es C-continua se sigue que M es constante en el intervalo $[C_{\hat{T}-}, C_{\hat{T}}]$ que contiene a $[T, C_{\hat{T}}]$. Vemos que entonces

$$M_{C_s \wedge T} = M_{C_{s \wedge \hat{T}}}.$$

Se concluye que $\hat{M}^{\hat{T}}$ es una martingala acotada y que por lo tanto $\hat{M}$ es una martingala local continua.

Si N y M son martingalas locales continuas y C-continuas entonces MN es C-continuo. Como los intervalos de constancia de M coinciden con los de $\langle M\rangle$, vemos que $\langle M\rangle$ es C-continua. Por el mismo argumento $\langle N\rangle$ es C-continua y por la fórmula de polarización, $\langle M, N\rangle$ es C-continua. Así, la martingala local continua $MN - \langle M, N\rangle$ es C-continua. Por la conclusión del párrafo anterior, vemos que $\hat{M}\hat{N} - \widehat{\langle M, N\rangle}$ es una martingala local continua. Se concluye que $\langle \hat{M}, \hat{N}\rangle = \widehat{\langle M, N\rangle}$.

Finalmente, sea $H \in L_2^{\mathrm{loc}}(M)$. Puesto que $\langle H \cdot M\rangle = H^2 \cdot \langle M\rangle$, vemos que $H \cdot M$ es constante en cualquier intervalo en el que $\langle M\rangle$ lo sea. Por lo tanto, $H \cdot M$ es C-continua. Concluimos que $\widehat{H \cdot M}$ es una $\hat{\mathscr{F}}_t$-martingala local continua con variación cuadrática $\widehat{H \cdot \langle M\rangle}$. Si (T_n) es una sucesión de tiempos de paro que crece casi seguramente a ∞ y tales que

$$\mathbb{E}\left(\int_0^{T_n} H^2 \, d\langle M\rangle\right) < \infty,$$

aplicamos el cambio de tiempo a la integral de Lebesgue-Stieltjes para obtener

$$\int_0^{T_n} H^2 \, d\langle M\rangle = \int_0^{C^{-1}_{T_n}} \hat{H}^2 \, d\langle \hat{M}\rangle.$$

Como $C_{T_n}^{-1}$ es un tiempo de paro respecto de $(\hat{\mathscr{F}}_t)$ (pues esta filtración es continua por la derecha y

$$\left\{C_{T_n}^{-1} < s\right\} = \{T_n < C_s\} \in \mathscr{F}_{C_s} = \hat{\mathscr{F}}_s)$$

y $C_{T_n}^{-1} \to \infty$ puesto que $C_t < \infty$ para toda $t \geq 0$, concluimos que $\hat{H} \in L_2^{\mathrm{loc}}(\hat{M})$. Por otra parte, vemos que

$$\begin{aligned}\langle \widehat{H \cdot M}, \hat{H} \cdot \hat{M} \rangle &= \hat{H} \cdot \langle \widehat{H \cdot M}, \hat{M} \rangle = \hat{H} \cdot \langle \widehat{H \cdot M, M} \rangle \\ &= \hat{H} \cdot \widehat{H \cdot \langle M \rangle} = \hat{H} \cdot \hat{H} \cdot \widehat{\langle M \rangle} = \hat{H}^2 \cdot \widehat{\langle M \rangle}.\end{aligned}$$

Vemos entonces que $\widehat{H \cdot M} - \hat{H} \cdot \hat{M}$ es una $(\hat{\mathscr{F}}_t)$-martingala local continua de variación cuadrática nula y se concluye que $\widehat{H \cdot M}$ y $\hat{H} \cdot \hat{M}$ son indistinguibles. □

Teorema 4.4 (Dambis-Dubins-Schwarz, [**DS65, Dam65**]). *Sea M una martingala local continua nula en cero y tal que $\langle M \rangle_\infty = \infty$. Entonces existe un movimiento browniano β tal que $M_t = \beta_{\langle M \rangle_t}$.*

Demostración. Puesto que $\langle M \rangle_\infty = \infty$, su inverso continuo por la derecha $\langle M \rangle^{-1}$ es finito casi-seguramente. Ya que M y $\langle M \rangle$ tienen los mismos intervalos de constancia y que $[\langle M \rangle_{t-}^{-1}, \langle M \rangle_t]$ es un intervalo de constancia de $\langle M \rangle$, se sigue que M es $\langle M \rangle^{-1}$-continua. Concluimos del Teorema 4.3 que β es una $\left(\hat{\mathscr{F}}_t\right)$-martingala local continua de variación cuadrática $\langle M \rangle \circ \langle M \rangle^{-1}$. Puesto que $\langle M \rangle$ es continuo, entonces $\langle M \rangle \circ \langle M \rangle^{-1}$ y por lo tanto β es un movimiento browniano como afirma el Teorema de Caracterización de Lévy. Además, se afirma que $M = \beta \circ \langle M \rangle$. En efecto, notemos que $\beta \circ \langle M \rangle = M \circ \langle M \rangle^{-1} \circ \langle M \rangle$. Por otra parte, como $[t, \langle M \rangle_{\langle M \rangle_t}^{-1}]$ es un intervalo de constancia para $\langle M \rangle$ (por ser no-decreciente), también lo es para M y por lo tanto:

$$M_t = M_{\langle M \rangle_{\langle M \rangle_t}^{-1}} = \beta_{\langle M \rangle_t}. \qquad \square$$

El teorema anterior es la clave para ver que ciertos procesos de interés son soluciones a ecuaciones diferenciales estocásticas. (Posteriormente, analizaremos a profundidad la ecuación diferencial estocástica satisfecha por la norma al cuadrado del movimiento browniano en dimensión δ.) También hay una versión del teorema anterior que no utiliza la hipótesis de que la variación cuadrática sea infinita. Sin embargo, el movimiento browniano se encuentra entonces en una extensión de nuestro espacio de probabilidad. Además, hay una versión multidimensional del teorema anterior conocido como Teorema de Knight.

Teorema 4.5 ([**Kni71**]). *Sean $M^1, \ldots, M^n$ martingalas locales continuas tales que $\langle M^i \rangle_\infty = \infty$ y $\langle M^i, M^j \rangle = 0$. Entonces existe un movimiento browniano n-dimensional $\beta = \left(\beta^1, \ldots, \beta^n\right)$ tal que $M^i = \beta^i \circ \langle M^i \rangle^{-1}$.*

Para comprender la complicación de este teorema sobre el de Dambis-Dubins-Schwarz, notemos que si para cada i definimos a $\beta^i = M^i \circ \langle M^i \rangle^{-1}$ y a $\mathscr{G}_t^i =$

$\mathscr{F}_{\langle M^i\rangle_t^{-1}}$, entonces β^i es un $\left(\mathscr{G}_t^i\right)$-movimiento browniano. Lo que no queda claro es si hay alguna filtración $(\mathscr{G}_t)$ para la cual β sea un $(\mathscr{G}_t)$-movimiento browniano o si al menos $\beta^1, \ldots, \beta^n$ son independientes. Sin embargo, hay un caso más sencillo en el que podemos dar una prueba similar a la del teorema de Dambis-Dubins-Schwarz:

Corolario 5 (Kunita-Watanabe, [**KW67**]). *Sean $M^1, \ldots, M^n$ martingalas locales continuas tales que $\langle M^i\rangle_\infty = \infty$, $\langle M^i, M^j\rangle = 0$ y $\langle M^i\rangle = \langle M^j\rangle$. Entonces existe un movimiento browniano n-dimensional $\beta = \left(\beta^1, \ldots, \beta^n\right)$ tal que $M^i = \beta^i \circ \langle M^i\rangle^{-1}$.*

Aunque la hipótesis parezca demasiado restrictiva, este teorema es útil pues ayuda a estudiar la liga entre el movimiento browniano en el plano y el análisis complejo.

Prueba del teorema de Knight (Teorema 4.5). Analizaremos la prueba de [**CY80**].

Definimos a

$$\beta^i = M \circ \langle M^i\rangle^{-1};$$

sabemos por el teorema DDS que cada β^i es un movimiento browniano. Ahora veremos que son independientes. Para esto consideraremos una partición $0 = t_0 < \cdots < t_m$ y mostraremos que

$$e^{-\frac{1}{2}\sum_{k=1}^n \sum_{j=1}^m \left(\lambda_j^k\right)^2 (t_j - t_{j-1})} = \mathbb{E}\left(e^{i\sum_{k=1}^n \sum_{j=1}^m \lambda_k^j \left(\beta_{t_j}^k - \beta_{t_{j-1}}^k\right)}\right).$$

Esto muestra que $\beta^1, \ldots, \beta^n$ son independientes.

Para realizar el cálculo de la esperanza denotemos por A^k a $\langle M^k\rangle$ y consideramos a

$$f_k(t) = \sum_{j=1}^m \lambda_j^k \mathbf{1}_{(t_{j-1}, t_j]},$$

así como a

$$N = \mathscr{E}\left(i\sum_{k=1}^n f_k \circ A^k \cdot M^k\right);$$

al utilizar la hipótesis sobre la covariación de M^i con M^j, vemos que

$$N_t = e^{i\sum_{k=1}^n \int_0^t f_k\left(A_s^k\right) dM_s^k - \frac{1}{2}\int_0^t f_k\left(A_s^k\right)^2 d\langle M^k\rangle_s}.$$

Al ser N una martingala local continua y acotada, es de hecho una martingala acotada y por lo tanto tiene un límite casi seguro y en L_1 que satisface

$$\mathbb{E}(N_\infty) = \mathbb{E}(N_0) = 1.$$

Por otra parte, al utilizar el teorema de cambio de variable para las integrales estocásticas (cf. Teorema 4.3) y de Riemann-Stieltjes, vemos que

$$\int_0^\infty f_k\left(A_s^k\right)\, dM_s^k = \int_0^\infty f_k(s)\, d\beta_s^k$$

y

$$\int_0^\infty f_k\left(A_s^k\right)\, d\langle M^k\rangle_s = \int_0^\infty f_k(s)\, ds.$$

Se concluye que

$$N_\infty = e^{i\sum_{k=1}^n \sum_{j=1}^m \lambda_k^j\left(\beta_{t_j}^k - \beta_{t_{j-1}}^k\right) + \frac{1}{2}\sum_{k=1}^n \sum_{j=1}^m \left(\lambda_j^k\right)^2 (t_j - t_{j-1})}$$

lo cual termina la prueba. □

4. La norma del movimiento browniano en $\mathbb{R}^d$

Sea $B = \left(B^1, \ldots, B^d\right)$ un movimiento browniano y definamos a

$$Z_t = \|\vec{x} + B_t\|^2 = \sum_{i=1}^d \left(x_i + B_t^i\right)^2.$$

Al aplicar la fórmula de Itô con la función $f(\vec{y}) = \|\vec{x} + \vec{y}\|^2$ a la semimartingala vectorial B, vemos que

$$Z_t = f(B_t) = x + \sum_{i=1}^d \int_0^t 2\left(x_i + B_s^i\right)\, dB_s^i + dt$$

donde $x = \|\vec{x}\|^2$. Definamos a

$$M_t = \sum_{i=1}^d \int_0^t 2\left(x_i + B_s^i\right)\, dB_s^i.$$

Entonces M es una martingala local continua con variación cuadrática

$$\sum_{i=1}^d \int_0^t 4Z_s\, ds.$$

Sea ahora $h : \mathbb{R} \to \mathbb{R}$ una función de clase C_2. Entonces

$$\begin{aligned} h(Z_t) &= h(x) + \int_0^t h'(Z_s)\, dZ_s + \frac{1}{2}\int_0^t h''(Z_s)\, 4Z_s\, ds \\ &= h(x) + \int_0^t h'(Z_s)\, dM_s + \int_0^t h'(Z_s)\,\delta + h''(Z_s)\, 2Z_s\, ds. \end{aligned}$$

Vemos entonces que si

$$2xh''(x) + \delta h'(x) = 0$$

entonces $h(Z)$ será una martingala local continua. Por otra parte, al probar funciones de la forma $h(x) = x^\alpha$, vemos que si $\delta \neq 2$ entonces

$$h(x) = x^{1-\delta/2}$$

satisface la ecuación diferencial anterior mientras que cuando $\delta = 2$, la función

$$h(x) = \log x$$

lo hace. Sean $0 < r < x < R$ y definamos a $T_{r,R}$ como la primera vez que B sale del anillo $\left\{\vec{x} : r < \|\vec{x}\|^2 < R\right\}$. En otras palabras, definamos a

$$T_r = \inf\{t \geq 0 : Z_t \leq r\}, \quad T^R = \inf\{t \geq 0 : Z_t \geq R\} \quad \text{y} \quad T_{r,R} = R_r \wedge T^R.$$

Notemos que $T_{r,R} < \infty$ casi seguramente puesto que las variables $\|B_{n+1} - B_n\|^2$ son independientes e idénticamente distribuidas y $\mathbb{P}\big(\|B_1\|^2 > 2R\big) > 0$. Por Borel-Cantelli, casi seguramente existe n tal que $\|B_{n+1} - B_n\|^2 > 2R$ y para dicha n forzosamente se sigue que $T_{r,R} \leq n$.

Puesto que $h\big(Z^{T_{r,R}}\big)$ es una martingala local continua acotada, es una martingala uniformemente integrable y por lo tanto

$$h(x) = \mathbb{E}(h(Z_0)) = \mathbb{E}\big(h\big(Z_{T_{r,R}}\big)\big) = h(r)\,p + h(R)\,(1-p) \quad \text{donde} \quad p = \mathbb{P}(T_r < T_R).$$

Se sigue que

$$\mathbb{P}(T_r < T_R) = \begin{cases} \frac{R^{1-\delta/2} - x^{1-\delta/2}}{R^{1-\delta/2} - r^{1-\delta/2}} & \delta \neq 2 \\ \\ \frac{\log R/x}{\log R/r} & \delta = 2 \end{cases}.$$

Puesto que las trayectorias de Z son continuas, se sigue que $T_R \to \infty$ conforme $R \to \infty$. Se deduce

$$\mathbb{P}(T_r < \infty) = \begin{cases} 1 & \delta \leq 2 \\ \left(\frac{x}{r}\right)^{1-\delta/2} & \delta > 2 \end{cases}.$$

Por otro lado, puesto que $T_r \to T_0$ conforme $r \to 0$ entonces vemos que

$$\mathbb{P}(T_0 < T_R) = \begin{cases} 0 & \delta \geq 2 \\ 1 - \left(\frac{x}{R}\right)^{1-\delta/2} & \delta < 2 \end{cases}.$$

Nótese que se ha utilizado $\delta < 2$ en vez de $\delta = 1$. Esto se sigue de que es posible definir a un proceso que actúe como la norma al cuadrado del movimiento browniano en dimensión δ para cualquier $\delta \geq 0$. En efecto, a continuación utilizaremos el teorema de Dambis-Dubins-Schwarz para verificar que cuando δ es entero no-negativo, entonces Z satisface una ecuación diferencial estocástica (parametrizada por δ). Se utilizará esta ecuación diferencial estocástica para darle sentido a Z cuando δ no es natural. Antes de eso, continuemos con algunas consecuencias de los cálculos que hemos hecho:

Corolario 6. *Sea B un movimiento browniano δ-dimensional que parte de cero. Si $\delta \geq 2$, B jamás regresa a cero. Si $\delta \leq 2$ entonces B regresa a cualquier vecindad de cero y el conjunto de puntos en que se encuentra en una vecindad de cero no es acotado. Si $\delta > 2$, B es transitorio.*

Demostración. Ya hemos probado que para $\delta \geq 2$ y $x \neq 0$ entonces $x + B$ jamás se anula. Apliquemos lo anterior al proceso $B_{\varepsilon+t}, t \geq 0$ donde $\varepsilon > 0$. Puesto que $B_\varepsilon \neq 0$ casi seguramente, al condicionar por B_ε, vemos que $B_{\varepsilon+t}, t \geq 0$ jamás se anula casi seguramente. Al ser válida esta conclusión para cualquier $\varepsilon > 0$, vemos que $B_t, t > 0$ jamás se anula.

También hemos visto que si $\delta \leq 2$ y $x \neq 0$ entonces $x + B$ regresa a cualquier vecindad (fija) de cero. Al aplicar esto al proceso $B_{t+n}, t \geq 0$, condicionalmente a B_n (que es casi seguramente distinto de cero), vemos que casi seguramente $B_{t+n}, t \geq 0$ regresa a cualquier vecindad fija V de cero. Así, para toda $n \geq 1$ existe $t_n \geq n$ tal que B_{t_n} pertenece a V y por lo tanto, el conjunto de visitas de B a V no es acotado.

Finalmente, si $\delta > 2$ y $\vec{x} \neq 0$ entonces $\|x + B_t\|^{2-\delta}$ es una martingala local no-negativa. Esto implica que se trata de una supermartingala no-negativa. Por lo tanto, converge casi seguramente a un límite finito, digamos ξ. Por el lema de Fatou y la autosimilitud del movimiento browniano vemos que

$$\mathbb{E}(\xi) \leq \liminf_{t\to\infty} \mathbb{E}\left(\frac{1}{\|x + B_t\|^{2-\delta}}\right) = \lim_{t\to\infty} \mathbb{E}\left(\frac{1}{\|x + \sqrt{t}B_1\|^{2-\delta}}\right) = 0. \qquad \square$$

Sean $\delta \geq 2$ y $\vec{x} \neq 0$. Puesto que $Z \neq 0$ casi seguramente, el proceso $1/2\sqrt{Z}$ es continuo, por lo que podemos definir al proceso β mediante

$$\beta = \frac{1}{2\sqrt{Z}} \cdot M.$$

El teorema de caracterización de Lévy nos dice que β es un movimiento browniano y por construcción

$$Z_t = \|x\|^2 + \int_0^t 2\sqrt{Z_s}\, d\beta_s + \delta t. \tag{4}$$

Por supuesto la ecuación diferencial anterior tiene sentido aún cuando δ no sea un entero positivo y esta es la manera en la que consideraremos al cuadrado de la norma del browniano δ-dimensional aún cuando δ no sea un entero. El único problema con la ecuación anterior es que no podemos utilizar el teorema de existencia y unicidad para ecuaciones diferenciales estocásticas puesto que el coeficiente de la ecuación no es Lipschitz.

Tomaremos un camino ad hoc para discutir la ecuación (4) . Comenzaremos por ver a Z como una solución débil, que es la que se puede construir de manera más sencilla. Sean $\sigma(x) = 2\sqrt{|x|}$ y $b = \delta \,\mathrm{Id}$ para $\delta \geq 0$.

Definición. Una **solución débil** a la ecuación diferencial estocástica con condición inicial $x \geq 0$, coeficiente de difusión σ y coeficiente de deriva b, denotada como $E(x, \delta)$, consta de un espacio de probabilidad y una filtración $(\mathscr{F}_t)$ que satisfaga las condiciones habituales en el que esté definido un $(\mathscr{F}_t)$-movimiento browniano β y un proceso continuo, no-negativo y $(\mathscr{F}_t)$-adaptado Z tal que

$$Z_t = x + \int_0^t \sigma(Z_s)\ d\beta_s + \int_0^t b(Z_s)\ ds.$$

Así, para el concepto de solución débil no está dado de antemano el movimiento browniano como el caso en que los coeficientes de deriva y difusión son Lipschitz. Recordemos que en este último caso, hemos visto que la solución es adaptada respecto del movimiento browniano que conduce a la ecuación, lo cual tampoco requerimos. Más adelante veremos ejemplos de ecuaciones diferenciales estocásticas para las cuales hay soluciones débiles pero tal que ninguna solución débil puede ser adaptada al browniano que la conduce. También veremos más adeltante por qué de hecho la ecuación diferencial estocástica que estamos considerando admite soluciones adaptadas.

Definición. Decimos que la ecuación $E(x, \delta)$ satisface unicidad débil si cada vez que Z y $\tilde{Z}$ sean soluciones débiles se tiene que Z y $\tilde{Z}$ tienen la misma distribución.

Equivalentemente, Z y $\tilde{Z}$ tienen las mismas distribuciones finito-dimensionales.

Teorema 4.6. *Para toda $x \geq 0$ y $\delta > 0$, hay existencia y unicidad débil para la ecuación $E(x, \delta)$.*

Por lo tanto, a cualquier proceso que sea solución débil a $E(x, \delta)$ le daremos el nombre de cuadrado de proceso de Bessel de dimensión $\delta \geq 0$ que comienza en x.

La idea de la prueba consiste en transformar la ecuación diferencial estocástica

$$Z_t = x + \int_0^t 2\sqrt{Z_s}\, d\beta_s + \delta t$$

para Z en la ecuación diferencial ordinaria

$$Z_t = x + B_{\int_0^t 4Z_s\, ds} + \delta t,$$

para $\int_0^t Z_s\, ds$ al utilizar el teorema de Dambis-Dubins-Schwarz. Como veremos, hay existencia y unicidad para dicha ecuación diferencial ordinaria y podemos escribir a $Z_t = F_t(B)$ para alguna funcional medible F_t. Así, las distribuciones finito-dimensionales de Z están determinadas por la funcional F_t aplicada al movimiento browniano B. Por otra parte, si Z resuelve la ecuación ordinaria entonces (modulo un detalle técnico que tiene que ver con los ceros de Z), se tiene que Z resuelve la ecuación diferencial estocástica respecto del movimiento browniano $\beta = 1/2\sqrt{Z} \cdot B$. Este método no nos permite fijar al browniano β desde el inicio, por lo que sólo produce soluciones débiles.

Demostración. Sea B un movimiento browniano. Consideremos, para $x \geq 0$ y $\delta > 0$ a la ecuación diferencial ordinaria

$$Z_t = x + B_{4\int_0^t Z_s\, ds} + \delta t, \tag{5}$$

que se puede poner en la forma $C'_t = x + B \circ 4C + \delta\,\mathrm{Id}$ donde

$$C_t = \int_0^t Z_s\, ds$$

y donde extendemos a B a $\mathbb{R}_-$ poniéndolo igual a cero.

LEMA 3. *Hay a lo más una solución a la ecuación diferencial* (5).

La prueba de este lema es muy adaptada a la forma especial de la ecuación y se dejará al final.

Ahora podemos mostrar la existencia; lo haremos al mostrar que el método de Euler converge. Antes que nada, notemos que por la ley fuerte de los grandes números $B_t/t \to 0$ conforme $t \to \infty$. Por lo tanto, existe una constante K tal que $|B_t| \leq Kt$ para toda $t \geq 0$. Ahora utilicemos el método de Euler: para cada $n \geq 0$, definamos a

$$C_0^n = 0 \quad C_{(k+1)/n}^n = C_{k/n}^n + \left[x + B_{4C_k^n} + \delta \frac{k}{n}\right]^+ \frac{1}{n}$$

y extendamos por interpolación en cada intervalo $[k/n, (k+1)/n]$. Se afirma que C^n es una sucesión de funciones continuas no-decrecientes y uniformemente acotada en compactos. En efecto, notemos que

$$C_{k/n}^n = \sum_{j<k} \left[x + B_{4C_j^n} + \delta \frac{j}{n}\right] \frac{1}{n} \leq \sum_{j<k} \left[x + 4KC_j^n + \delta \frac{j}{n}\right] \frac{1}{n}.$$

por lo que si $k/n \leq t$ tenemos que

$$C_t^n \leq \int_0^t x + 4KC_s^n \delta t \, ds.$$

Por la desigualdad de Gronwall, vemos que

$$C_t^n \leq t(x + \delta t)e^{4Ks}$$

si $s \in [0, t]$. Puesto que C^n es uniformemente acotada en $[0, t]$ entonces también es uniformemente equicontinua. En efecto, si Z^n es la derivada lateral derecha de C^n, se sigue de la definición por recurrencia de C^n que

$$|Z_t^n| \leq \left[x + B_{C_{k/n}^n} + t\right]^+ \quad t \in [k/n, (k+1)/n].$$

Por el teorema de Arzelà-Ascoli, existe un proceso C con trayectorias continuas y una subsucesión n_k tal que C^{n_k} converge uniformemente a C. La relación de recurrencia para C^n se puede replantear en la forma integral:

$$C_t^n = \int_0^t \left[x + B_{C_{\lfloor ns \rfloor/n}^n} + \delta \lfloor ns \rfloor/n\right]^+ \, ds$$

y al pasar al límite vemos que

$$C_t = \int_0^t [x + B_{C_s} + \delta s]^+ \, ds.$$

Así, vemos que C es no decreciente y derivable, por lo que su derivada es no-negativa, podemos omitir los valores absolutos y concluir que C satisface (5). Por unicidad para las soluciones a la ecuación (5), todos los límites subsucesionales de C^n son iguales y por lo tanto se concluye que el método de Euler es convergente a la única solución a la ecuación (5).

Si C es tal que

$$C_t = \int_0^t x + B_{4C_s}\, ds + \delta t.$$

La derivada de C, digamos Z, es la única solución a (5). Además, puesto que C es estrictamente creciente, entonces Z es no-negativo. Sea $\tilde{B}$ un movimiento browniano independiente de B y consideremos al proceso

$$\beta = \frac{\mathbf{1}_{Z\neq 0}}{2\sqrt{Z}} \cdot B \circ 4C + \mathbf{1}_{Z=0} \cdot \tilde{B}.$$

Entonces

$$\langle \beta \rangle = \mathbf{1}_{Z\neq 0} \frac{1}{4Z} \cdot 4C + \mathbf{1}_{Z=0} \cdot \mathrm{Id} = \mathrm{Id}$$

y por el teorema de caracterización de Lévy β es un movimiento browniano. Además, por definición, tenemos que

$$Z = x + 2\sqrt{Z} \cdot \beta + \delta\, \mathrm{Id},$$

por lo que Z satisface $E(x, \delta)$ y por lo tanto, hay existencia débil para $E(x, \delta)$.

Para mostrar unicidad débil para $E(x, \delta)$, sea Z una solución a dicha ecuación conducida por el browniano β:

$$Z_t = x + \int_0^t 2\sqrt{Z_s}\, d\beta_s + \delta t.$$

Por el teorema de Dambis-Dubins-Schwarz, existe un movimiento Browniano B (en una extensión de nuestro espacio de probabilidad) tal que

$$Z_t = x + B_{\int_0^t 4Z_s\, ds} + \delta t.$$

Sea $C_t = \int_0^t Z_s\, ds$. Puesto que $C_0 = 0$ y C es continuamente diferenciable, la distribución de C determina a la de Z y viceversa. Por otra parte, por la convergencia del método de Euler, C se puede ver como el límite conforme $n \to \infty$ de la sucesión de procesos constantes por pedazos y adaptados C^n dados por

$$C^n_0 = 0 \quad C^n_{(k+1)/n} = C^n_{(k+1)/n} + \left[x + B_{4C^n_k} + \delta \frac{k}{n} \right] \frac{1}{n}.$$

La recurrencia, en la que substituimos a B por cualquier función continua f es una funcional $F_n(f)$ tal que $F_n(B) = C^n$. Sea F la función medible $\limsup_n F_n$. Hemos visto que $F_n(B) \to F(B)$ uniformemente en compactos casi seguramente por lo que, para cualquier browniano B, $F(B)$ es la única solución a la ecuación (5). Si Z^1 y Z^2 son dos soluciones a 4 conducidas por brownianos β^1 y β^2 entonces existen dos movimientos brownianos B^1 y B^2 tales que $Z^i = F(B^i)$. Puesto que $B^1 \stackrel{d}{=} B^2$ vemos que $Z^1 \stackrel{d}{=} Z^2$. □

Dejamos pendiente la prueba de un lema ad hoc para nuestra ecuación.

Prueba del Lema 3. Primero mostremos que Z es no-negativa. En efecto, si existe t^* tal que $Z_{t*} < 0$ entonces, por la continuidad de Z, existen $t \geq 0$ y $\varepsilon > 0$ tales que $Z_t = 0$ y $Z < 0$ en $(t, t+\varepsilon)$. Se afirma que entonces $t > 0$ y $C_t > 0$. Si $t = 0$ entonces $Z_0 = 0$ (y por lo tanto $x = 0$) y $C < 0$ en $(0, \varepsilon)$ por lo cual $Z = \delta \operatorname{Id} \geq 0$ en $(0, \varepsilon)$, una contradicción. Por otra parte, si $C_t \leq 0$ entonces $C < 0$ en $(t, t+\varepsilon)$, por lo que $x = 0$ y $Z_s = \delta(s-t) \geq 0$ en $(t, t+\varepsilon)$, una contradicción. Ahora veremos que existen $t_1 < t < t_2 < t+\varepsilon$ tales que $Z_{t_1} > 0$ y $C_{t_1} = C_{t_2}$, lo que implica la contradicción

$$0 < Z_{t_1} = x + B_{4C_{t_1}} + \delta t_1 = x + B_{4C_{t_2}} + \delta t_1 < x + B_{4C_{t_2}} + \delta t_2 < 0$$

y nos muestra que Z no puede tomar valores estrictamente negativos. La existencia de t_1 y t_2 se sigue del siguiente argumento: puesto que $t > 0$ y $C_t > 0$, podemos suponer sin pérdida de generalidad que $C > 0$ en $(t, t+\varepsilon)$. Sea τ_1 la última vez anterior a t que C está por debajo de $C_{t+\varepsilon}$ (por lo que $\tau_1 < t$) y τ_2 la primera vez después de τ_1 tal que C iguala a C_t, por lo que $\tau_2 \leq t$. Puesto que $C_{\tau_2} - C_{\tau_1} = C_t - C_{t+\varepsilon} > 0$, existe $t_1 \in (\tau_1, \tau_2)$ tal que $Z_{t_1} > 0$. Pero por definición de τ_i, $C_{t_1} = C_{t_2}$ para alguna $t_2 \in (t, t+\varepsilon)$.

Ahora mostraremos que C es estríctamente creciente. En efecto, la otra opción es que C tenga un intervalo de constancia $[s, t]$ con $s < t$. Esto nos lleva a la contradicción

$$0 = Z_{(t+s)/2} = x + B_{4C_{(t+s)/2}} + \delta(t+s)/2 > x + B_{4C_s} + \delta s = Z_s \geq 0$$

puesto que $Z \geq 0$.

Ahora mostremos unicidad para la ecuación diferencial ordinaria. Sean Z y $\tilde{Z}$ dos soluciones con integrales C y $\tilde{C}$. Puesto que C y $\tilde{C}$ son estrictamente crecientes, podemos considerar a sus inversas I y $\tilde{I}$. Si $C \neq \tilde{C}$ entonces $I \neq \tilde{I}$. Supongamos que existe x_1 tal que $I_{x_1} < \tilde{I}_{x_1}$. Sea x_0 la última vez antes de x_1 tal que I supera a $\tilde{I}$:

$$x_0 = \sup \left\{ x \leq x_1 : I_x \geq \tilde{I}_x \right\}.$$

Puesto que C es continuamente diferenciable y estrictamente creciente, lo mismo le pasa a I e

$$I_y = \int_0^y \frac{1}{x + B_{4s} + \delta I_s} \, ds.$$

Debe de existir $z \in [x_0, x_1]$ tal que I'_z e $\tilde{I}'_z$ existen y $I'_z < \tilde{I}'_z$. Para este valor de z tenemos que

$$x + B_{4z} = \frac{1}{I'_z} - \delta I_z > \frac{1}{\tilde{I}_z} - \delta \tilde{I}_z = x + B_{4z}.$$

Por lo tanto $I = \tilde{I}$ y $C = \tilde{C}$. $\square$

Supongamos que Z^1 y Z^2 resuelven las ecuaciones

$$Z_t^i = x_t + \int_0^t 2\sqrt{Z_s^i}\, d\beta_s^i + \delta_i t$$

donde β^1 y β^2 son movimientos brownianos independientes y Z^i es adaptado respecto de la filtración completada de β^i. Entonces el proceso $Z^1 + Z^2$ satisface:

$$Z_t^1 + Z_t^2 = (x_1 + x_2) + 2\int_0^t Z_s^1\, d\beta_s^1 + Z_s^2\, d\beta_s^2 + (\delta_1 + \delta_2)\, t.$$

Sea β^3 un movimiento browniano independiente de β^1 y β^2 (que jugará un rol técnico pero irrelevante). Definamos a

$$\beta_t = \int_0^t \frac{Z_s^1}{\sqrt{Z_s}} \mathbf{1}_{Z_s \neq 0}\, d\beta_s^1 + \int_0^t \frac{Z_s^2}{\sqrt{Z_s}} \mathbf{1}_{Z_s \neq 0}\, d\beta_s^2 + \int_0^t \mathbf{1}_{Z_s = 0}\, d\beta_s^3.$$

Al calcular la variación cuadrática de β y utilizar el teorema de Dambis-Dubins-Schwarz vemos β es un movimiento browniano. Por lo tanto, Z satisface la ecuación diferencial estocástica

$$Z_t = (x_1 + x_2) + 2\int_0^t \sqrt{Z_s}\, d\beta_s + (\delta_1 + \delta_2)\, t.$$

Se concluye que Z es el cuadrado de un proceso de Bessel de dimensión $\delta_1 + \delta_2$ que comienza en $x_1 + x_2$. A esta propiedad se le conoce como aditividad de los cuadrados de procesos de Bessel. Fue introducida por Shiga y Watanabe en [**SW73**].

5. El movimiento browniano complejo

En esta sección estudiaremos algunos resultados sobre el movimiento browniano en dimensión 2 y los utilizaremos para dar demostraciones probabilísticas de algunos resultados de variable compleja.

Definición. Una **martingala local compleja** es un proceso estocástico $Z = X + iY$ tal que tanto X como Y son martingalas locales reales.

Un ejemplo de esto sería el movimiento browniano complejo $B = B^1 + iB^2$ donde B^1 y B^2 son movimientos brownianos independientes.

Definimos la **variación cuadrática** de una martingala local compleja $Z = X + iY$ mediante la fórmula

$$\langle Z \rangle = \langle X \rangle - \langle Y \rangle + i2\langle X, Y \rangle.$$

Esta definición está basada en la bilinealidad de la covariación e implica que

$$Z^2 - \langle Z \rangle = \left(X^2 - Y^2 - \langle X \rangle + \langle Y \rangle\right) + i2\left(XY - \langle X, Y \rangle\right)$$

sea una martingala local compleja. Sin embargo, notemos que si utilizamos la definición de producto complejo, entonces

$$\langle Z \rangle = \mathbb{P} - \lim_{|\Delta| \to 0} \sum_{t_i \in \Delta} \left(Z_{t_i} - Z_{t_{i-1}} \right)^2 .$$

De igual manera, $H = H^r + iH^i$ es un proceso complejo tal que sus partes reales e imaginarias son localmente acotadas y $Z = X + iY$ es una martingala local compleja, definimos

$$H \cdot Z = \left[H^i \cdot X - H^i \cdot Y \right] + i \left[H^i \cdot X + H^r \cdot Y \right]$$

y, si H es continuo por la izquierda, se tiene la aproximación

$$(H \cdot Z)_t = \mathbb{P} - \lim_{|\Delta| \to 0} \sum_{\Delta} H_{t_{i-1}} \left(Z_{t_i} - Z_{t_{i-1}} \right) .$$

Con esta definición, vemos que si $H \cdot Z$ es de nuevo una martingala local compleja.

Ejercicio 4.1. Pruebe de la definción que

$$\langle H \cdot Z \rangle = H^2 \cdot \langle Z \rangle.$$

El movimiento browniano no sólo es una martingala local compleja sino que tiene una propiedad adicional.

Definición. Sea $Z = X + iY$ una martingala local compleja. Decimos que Z es una **martingala local conforme** si Z^2 es una martingala local compleja.

Evidentemente, esto es equivalente a que $\langle Z \rangle = 0$ o inclusive a que $\langle X \rangle = \langle Y \rangle$ y $\langle X, Y \rangle = 0$. Entonces, vemos que el movimiento browniano complejo es una martingala local conforme.

El siguiente teorema representa un primer paso para probar a la invariancia conforme de las trayectorias brownianas, que ha sido objeto de mucho estudio y progreso reciente. Recordemos que una función $f : \mathbb{C} \to \mathbb{C}$ es entera si es derivable en el sentido complejo. Esto es, si para todo $z \in \mathbb{C}$ existe un complejo $Df(z)$ tal que

$$Df(z) = \lim_{\substack{h \to 0 \\ h \in \mathbb{C}}} \frac{f(z+h) - f(z)}{h}.$$

Teorema 4.7. *Sea $B = B^1 + iB^2$ un movimiento browniano complejo y f una función entera no-constante. Entonces $f(B)$ es una martingala local conforme, $\langle \operatorname{Re} f(B) \rangle_\infty = \infty$ casi seguramente y existe un movimiento browniano complejo β tal que*

$$f(B_t) = f(0) + \beta_{\langle \operatorname{Re} f(B) \rangle_t}.$$

Demostración. Podemos utilizar la fórmula de Itô a $f(B)$ para escribir

$$f(B) = f(B_0) + Df(B) \cdot B,$$

puesto que

$$D^2 f(B) \cdot \langle B \rangle = 0.$$

Ejercicio 4.2. Probar la fórmula anterior al utilizar la definición de integral estocástica respecto de una martingala local compleja y la fórmula de Itô en el caso real. Sugerencia: Las ecuaciones de Cauchy-Riemann serán útiles.

Intuitivamente, dicha fórmula debe valerse al utilizar Taylor a orden 2 como en el caso del movimiento browniano y utilizando el hecho de que la variación cuadrática del browniano complejo se anula. Ahora, puesto que $\langle Df(B) \cdot B \rangle = Df(B)^2 \cdot \langle B \rangle = 0$, vemos que $f(B)$ es una martingala local conforme.

Para aplicar el teorema DDS y construir al movimiento browniano β, primero mostraremos que la variación cuadrática de $\operatorname{Re} f(B)$ tiende a infinito en infinito. Notemos que

$$\langle \operatorname{Re} f(B) \rangle = \| Df(B) \|^2 \cdot \operatorname{Id}.$$

Puesto que Df no es constante, existe $z \in \mathbb{C}$ y $r > 0$ tal que Df no se anula en $\overline{B_r(z)}$; sea $\varepsilon > 0$ el mínimo de $\|Df\|^2$ en dicho compacto. Vemos entonces que

$$\int_0^\infty \| Df(B_s) \|^2 \, ds \geq \varepsilon \int_0^\infty \mathbf{1}_{B_s \in \overline{B_r(z)}} \, ds.$$

Ahora generalizaremos un poco nuestro resultado sobre recurrencia en el sentido de las vecindades del movimiento browniano plano para mostrar que la integral del lado derecho es infinita, lo cual nos dice que $\langle \operatorname{Re} f(B) \rangle_\infty = \infty$ casi seguramente. En efecto, sean

$$T_0 = \inf \{ t \geq 0 : \| B_t - z \| \geq r \} \qquad T_{2n+1} = \inf \{ t \geq T_{2n} : \| B_t - z \| \leq r/2 \}$$

y

$$T_{2n+2} = \inf \{ t \geq T_{2n+1} : \| B_t - z \| \geq r \}.$$

Hemos visto que el movimiento browniano no es acotado casi seguramente. Por lo tanto, $T_0 < \infty$ y $T_{2n+1} < \infty$ si $T_{2n} < \infty$ (al utilizar la propiedad de Markov fuerte). Por la recurrencia en el sentido de las vecindades hemos visto que si $T_{2n} < \infty$ entonces $T_{2n+1} < \infty$. Por lo tanto $T_n < \infty$ para toda n casi seguramente. La propiedad de Markov fuerte del movimiento browniano y su invarianza ante rotaciones nos dice que las variables $B_{T_n} - z$ son uniformes en $\{ \| \tilde{z} - z \| = r/2 \}$ para n par y en $\{ \| \tilde{z} - z \| = r \}$ para n impar. Por la propiedad de Markov fuerte de nuevo, vemos que las variables

$$\int_{T_{2n+1}}^{T_{2n+2}} \mathbf{1}_{r \leq \| B_t - z \| \leq R} \, dt$$

son iid y estrictamente positivas. Por lo tanto,

$$\int_0^\infty \mathbf{1}_{r/2\leq\|B_t\|\leq r}\,dt = \infty$$

casi seguramente. En resumen,

$$\int_0^\infty \mathbf{1}_{\|B_t-z\|\leq r} = \infty$$

casi seguramente.

Podemos entonces utilizar el teorema de Kunita-Watanabe para deducir que existe un movimiento browniano β tal que

$$f(B) = f(0) + \beta \circ \langle \operatorname{Re} f(B)\rangle. \qquad \square$$

Corolario 7 (Teorema de Liouville). *Toda función entera y acotada es constante.*

Demostración. Sea f una función entera y acotada. Si f no fuera constante entonces existiría un movimiento browniano β tal que

$$f(B_t) = \beta_{\langle \operatorname{Re} f(B)\rangle_t}.$$

Esto es una contradicción puesto que el movimiento browniano β no es acotado y puesto que f no es constante hemos visto que $\langle \operatorname{Re} f(B)\rangle_\infty = \infty$ y por lo tanto $\beta \circ \langle \operatorname{Re} f(B)\rangle$ tampoco es acotado. $\square$

Corolario 8. *Sea p un polinomio complejo. Si p no tiene raíces entonces es constante.*

Demostración. Es bien sabido cómo probar este resultado a partir del teorema de Liouville. Si p no admite raíces. Entonces $1/p$ es una función entera y acotada y por lo tanto constante, por lo que a su vez p es constante. $\square$

A su vez, el resultado anterior permite fácilmente probar que un polinomio complejo de grado n admite n raíces complejas.

También es posible probar el siguiente teorema al utilizar al movimiento browniano:

Teorema 4.8 (Teorema pequeño de Picard). *Si f es entera y no es constante entonces su imagen omite a lo más un elemento de $\mathbb{C}$.*

Es importante permitir que se pueda omitir algún valor, como lo muestra la función exponencial, que jamás se anula. Para la prueba de este teorema y algunos otros resultados de variable compleja, ver [**Dav79**].

6. Movimiento browniano y funciones armónicas

En esta sección, veremos como el movimiento browniano nos permite resolver la ecuación de Poisson. El lector puede consultar un desarrollo más a profundidad de estos temas en [**PS78**] y [**Doo84**].

Sea $\delta \in \mathbb{Z}_+$ y consideremos un abierto $D \subset \mathbb{R}^\delta$ con cerradura $\overline{D}$ y frontera $\partial(D)$. Consideremos además una condición de frontera de Dirichlet $f : \partial(D) \to \mathbb{R}$ continua y un término que representa la fuente de calor externa $g : D \to \mathbb{R}$. Una solución a la ecuación de Poisson es una función continua $u : \overline{D} \to \mathbb{R}$ de clase C_2 en D y tal que

$$\begin{cases} \Delta u(x) = -g(x) & x \in D \\ u(x) = f(x) & x \in \partial(D) \end{cases}.$$

Si $g = 0$, la ecuación de Poisson resultante se denomina ecuación de Laplace. Sea $B = \left(B^1, \ldots, B^\delta\right)$ un movimiento browniano; utilizaremos a B para dar un resultado de unicidad para la ecuación de Poisson.

Teorema 4.9. *Supongamos que D, f y g son acotadas y que u es solución a la ecuación de Poisson. Sea*

$$S = \inf\left\{t \geq 0 : B_t \notin \overline{D}\right\}.$$

Entonces

$$u(x) = \mathbb{E}_x\left(f(B_S) + \int_0^S g(B_s)\, ds\right)$$

para toda $x \in D$.

Demostración. Sea $M_t = u\left(B_t^S\right) + \int_0^{t\wedge S} g(B_s)\, ds$. Al utilizar la fórmula de Itô, vemos que

$$\begin{aligned} M_t &= u(x) + \sum_i \int_0^{t\wedge S} D_i u(B_s)\, dB_s^i + \frac{1}{2}\int_0^{t\wedge S} \Delta u(B_s)\, ds + \int_0^{t\wedge S} g(B_s)\, ds \\ &= u(x) + \sum_i \int_0^{t\wedge S} D_i u(B_s)\, dB_s^i, \end{aligned}$$

donde la última igualdad se deduce pues u satisface la ecuación de Poisson.

Así, M es una martingala local. Puesto que u es continua y $\overline{D}$ es acotado, se sigue que u es acotada. Por lo tanto M es una martingala acotada. Además, al ser D acotado, se sigue que $S < \infty$ casi seguramente y por lo tanto

$$M_t \to u(B_S) + \int_0^S g(B_s)\, ds = f(B_S) + \int_0^S g(B_s)\, ds,$$

casi seguramente y en L_1. Al aplicar muestreo opcional, se sigue que

$$\mathbb{E}(M_S) = \mathbb{E}(M_0)$$

lo cual significa que

$$u(x) = \mathbb{E}_x\left(f(B_S) + \int_0^S g(B_s)\, ds \right). \qquad \square$$

El problema de existencia de soluciones a las ecuaciones de Poisson o Laplace es más delicado. Nos enfocaremos en el caso de la ecuación de Laplace.

Una función u de clase C_2 en D que satisface $\Delta u = 0$ se denomina **armónica**. Ahora veremos que una función armónica satisface la propiedad del promedio. Sea μ_r la distribución en $\left\{x \in \mathbb{R}^d : \|x\| = r\right\}$ dada por

$$\mu_r(A) = \mathbb{P}(B_{S_r} \in A) \quad \text{donde} \quad S_r = \inf\left\{t \geq 0 : \|B_t\| \geq r\right\}.$$

Puesto que B es invariante ante rotaciones, lo mismo le sucede a μ_r. Es por esto que es natural pensar en μ_r como la medida uniforme sobre la esfera de radio r.

Definición. Una función $u : D \to \mathbb{R}$ satisface la **propiedad del promedio** si para todo $x \in D$ tal que $\{\tilde{x} : \|\tilde{x} - x\| \leq r\} \subset D$ se tiene que

$$u(x) = \int u(x + \tilde{x})\, \mu_r(\tilde{x}).$$

Proposición 4.2. *Si u es armónica en D entonces satisface la propiedad del promedio en D.*

Demostración. Por supuesto, esta es una aplicación del teorema de la divergencia. Sin embargo, también es instructivo probarlo con el movimiento browniano.

En efecto, al aplicar la fórmula de Itô, vemos que $u\left(B^S\right)$ es una martingala acotada bajo $\mathbb{P}_x$ si S es el tiempo de salida de la bola de radio r con centro en x. Por muestreo opcional, vemos que

$$u(x) = \mathbb{E}_x(u(B_S)) = \int u(x + \tilde{x})\ \mu_r(d\tilde{x}). \qquad \square$$

La forma en la que usualmente se prueba unicidad para la ecuación de Laplace consiste en probar el principio del máximo, que es una consecuencia inmediata de la propiedad del promedio.

Corolario 9. *Si D es conexo, u es armónica en D y u se maximiza en D entonces u es constante en D.*

Demostración. Sea $M = \sup_{x \in D} u(x)$ y $D_M = \{x \in D : u(x) = M\}$. Claramente M es cerrado. Ahora utilizaremos la propiedad del promedio para verificar que D_M también es abierto. Así, concluiremos que $D_M = \emptyset$ ó $D_M = D$ pues D es conexo. La hipótesis de que u se maximice en D implica que $D_M \neq \emptyset$ lo cual a su vez nos dice que $D = D_M$.

Sea $x \in D_M$ y $r > 0$ tal que $B_\delta(x) \subset D$. Al utilizar la propiedad del promedio se ve que

$$u(x) = \frac{\int_{B_\delta(0)} u(x+\tilde{x})\; dx}{\int_{B_r(0)}\; dx}.$$

La fórmula anterior nos dice que $u = u(x)$ en $B_\delta(x)$, por lo que $B_\delta(x) \subset D_M$. $\square$

Para continuar con el análisis de existencia de soluciones a la ecuación de Laplace, se utilizará la siguiente proposición.

PROPOSICIÓN 4.3. *Si $u : D \to \mathbb{R}$ satisface la propiedad del promedio entonces es infinitamente diferenciable y armónica.*

La prueba es analítica y por lo tanto sólo se esbozará.

DEMOSTRACIÓN. La idea es que la convolución de u con una función radial coincide con u al utilizar el promedio y al escoger la función radial de tal manera que sea infinitamente diferenciable y de soporte compacto, se puede probar que u también es infinitamente diferenciable.

Para verificar que u es armónica se hace una expansión de Taylor a orden 2, se utiliza la simetría ante reflexiones de μ_r y la propiedad del promedio para ver que

$$0 = \sum_i \frac{\partial^2 u}{\partial x_i^2} \int_{\partial(B_\varepsilon(x))} \tilde{x}_i^2\, \mu_r(d\tilde{x}) + o\big(\varepsilon^2\big)\,,$$

y por intercambiabilidad

$$\int_{\partial(B_\varepsilon(x))} \tilde{x}_i^2\, \mu_r(d\tilde{x}) = \frac{1}{\delta} \sum_i \int_{\partial(B_\varepsilon(x))} \tilde{x}_i^2\, \mu_r(d\tilde{x}) = \frac{\varepsilon^2}{\delta}.$$

$\square$

Para resolver la ecuación de Laplace, supongamos que el tiempo de salida S de la región D es finito casi seguramente cuando el browniano parte de x para toda $x \in D$. (Esto ocurrirá si D es acotado.) Asumamos también que f es acotada, lo cual nos permite definir a

$$u(x) = \mathbb{E}_x\big(f\big(B^S\big)\big)\,.$$

Claramente u satisface la propiedad del promedio y por lo tanto es armónica. Sin embargo, no es inmediato ver que u sea continua y de hecho esto podría fallar. Para obtener la continuidad, se necesita imponer una condición de regularidad a la frontera de D:

DEFINICIÓN. Decimos que $x \in \partial(D)$ es un **punto regular** para D si

$$\mathbb{P}_x(S = 0) = 1.$$

TEOREMA 4.10. *Si D es acotado, f es continua y acotada y todo punto de $\partial(D)$ es regular para D entonces la función*

$$u(x) = \mathbb{E}_x\big(f\big(B^S\big)\big)$$

es la única solución al problema de Dirichlet para la ecuación de Laplace en D con condición a la frontera f.

Por supuesto, es necesario un criterio para la regularidad de la frontera. Sin embargo, dado lo que conocemos sobre el movimiento browniano no es demasiado difícil probar lo siguiente: si $x \in \partial(D)$ es tal que existen $y \in \mathbb{R}^d \setminus \{0\}$, $r > 0$ y $\theta \in (0, \pi)$ tal que $\left\{z \in \mathbb{R}^d : |(z-x) \cdot y| \geq \|z-x\| \|y\| \cos(\theta), \|z\| \leq r\right\} \subset \mathbb{R}^d \setminus D$ entonces z es regular. La última condición es la llamada **condición de cono de Zaremba**.

7. La fórmula de Feynman-Kac

La fórmula de Feynman-Kac es introducida por Kac en [**Kac49**] para calcular la distribución F_t de la variable aleatoria

$$A_t = \int_0^t v(B_s)\, ds$$

donde $v \geq 0$ satisface ciertas condiciones y B es un movimiento browniano. Un caso particular es cuando $v = \mathbf{1}_{(0,\infty)}$, en cuyo caso la distribución de A_t/t había sido encontrada por Lévy en [**Lév39**] y coincide con la llamada distribución arcoseno, que es la distribución de una variable Beta de parámetros $1/2$ y $1/2$. Las investigaciones de Kac siguen a unas anteriores de Erdös y Kac publicadas en [**EK46**] y [**EK47**] en la que consideran teoremas límites para funcionales de caminatas aleatorias y ven que en ciertos casos no dependen de la distribución de salto de la caminata aleatoria. De hecho, el punto de vista de Kac para encontrar la distribución de A_t es discretizar a A_t, encontrar una ecuación en diferencias para calcular la distribución de la aproximación, resolverla y pasar al límite. El nombre de Feynman aparece en la fórmula puesto Kac argumenta que su método está influenciado fuertemente por la derivación de Feynman de la ecuación de Shrödinger. En [**Sim05**] se pueden consultar aplicaciones de la medida de Wiener a la física cuántica con una discusión sobre la fórmula de Feynman-Kac.

El resultado de Kac es el siguiente: argumenta que si tomamos la doble transformada de Laplace de la distribución de A_t al definir

$$z(s, \alpha) = \int_0^\infty \int_0^\infty e^{-st-\alpha x} F_t(x)\, dtdx$$

entonces

$$z(s, \alpha) = \int_{-\infty}^\infty \psi(x, s, \alpha)\, dx$$

donde ψ resuelve la ecuación

$$\frac{1}{2} \frac{\partial^2 \psi}{\partial x} - (s + uv(x))\, \psi = 0$$

con las condiciones

$$\lim_{x\to\pm\infty}\psi(x,s,\alpha)=0,\qquad \left|\frac{\partial\psi}{\partial x}\right| \text{ es acotada y } \quad \psi'(0+)-\psi'(0-)=2.$$

Nosotros nos concentraremos en la derivación de la distribución de A_t cuando $v=\mathbf{1}_{(0,\infty)}$.

La formulación moderna de la fórmula de Feynman-Kac nos presenta una liga entre ecuaciones diferenciales parabólicas y difusiones. En efecto, nos afirma (en el caso unidimensional) que si existe una solución $u(t,x)$ a la ecuación

$$\frac{\partial u}{\partial t}+b\frac{\partial u}{\partial x}+\sigma^2\frac{\partial^2 u}{\partial x^2}+f=vu$$

para $u:[0,T]\times\mathbb{R}\to\mathbb{R}$ donde b,σ,f y v dependen de t y de x y se satisface la condición terminal

$$u(x,T)=\psi(x)$$

entonces u está dada por la fórmula

$$u(t,x)=\mathbb{E}\left(\int_t^T e^{-\int_t^{t_1} v\left(X_{t_2}\right)\,dt_2}f(t_1,X_{t_1})\ dt_1+e^{-\int_t^T v\left(X_{t_1}\right)\,dt_1}\psi(X_T)\right),$$

donde se asume que X satisface la ecuación diferencial estocástica

$$X_t=x+\int_0^t\sigma(s,X_s)\ dB_s+\int_0^t b(x,X_s)\ ds.$$

En particular, lo anterior representa un resultado de unicidad bajo el supuesto probabilístico de existencia débil a la ecuación diferencial estocástica.

Ahora veremos cómo probar dichos resultados, enfocándonos en casos particulares que muestren las ideas principales.

Comencemos con la liga entre el movimiento browniano y la ecuación de calor. Sea f medible y acotada y definamos

$$u(t,x)=\mathbb{E}_x(f(B_t)).$$

Notemos que $u(0,x)=f(x)$. Por otra parte, la propiedad de Markov nos dice que

$$u(t-s,B_s)=\mathbb{E}_x(\,f(B_s)\mid\mathscr{F}_s)\,,$$

por lo que $u(t-s,B_s)$ es una martingala en $[0,t]$. Por otra parte, si u fuera de clase C_2, podríamos utilizar la fórmula de Itô y concluir que

$$\begin{aligned}u(t-s,B_s)=&u(t,B_0)+\sum_{i=1}^{\delta}\int_0^s D_{1+i}u(t-r,B_r)\ dB_r^i\\&-\int_0^s D_1u(t-r,B_r)\ dr+\frac{1}{2}\int_0^s\Delta u(t-r,B_r)\ dr,\end{aligned}$$

donde el Laplaciano considera sólo a las últimas δ variables. Vemos por lo tanto que una condición natural para que al menos $u(t-s, B_s)$ sea una martingala local es que u satisfaga la ecuación de calor

$$\frac{\partial u}{\partial t} - \frac{1}{2}\Delta u = 0.$$

El siguiente resultado no es por lo tanto muy sorprendente.

Proposición 4.4. *Si u es continua en $[0,\infty)\times\mathbb{R}^\delta$, de clase C_2 en $(0,\infty)\times\mathbb{R}^\delta$ y satisface el problema de Cauchy*

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{1}{2}\Delta u = 0 \\ u(0,x) = f(x) \end{cases} \tag{6}$$

para alguna función continua y acotada f, entonces

$$u(t,x) = \mathbb{E}_x(f(B_t)).$$

Demostración. Probemos primero, mediante un argumento analítico, que u es acotada. En efecto, se afirma que para toda $\delta > 0$ y $M > 0$,

$$\max_{\delta\le t\le T, \|x\|\le M} u(t,x) \le \max_{\|x\|\le M} u(\delta, x).$$

En efecto, sean $\varepsilon > 0$ y $v(t,x) = u(t,x) - \varepsilon t$ y supongamos que v se maximiza en el interior de $[\delta, t]\times\{\|x\|\le M\}$, digamos en (t^*, x^*). Notemos primero que

$$\frac{\partial v}{\partial t} - \Delta v = \frac{\partial u}{\partial t} - \varepsilon - \Delta u = -\varepsilon.$$

Por otra parte, puesto que v se maximiza en (t^*, x^*), vemos que

$$\frac{\partial v}{\partial t}(t^*, x^*) \ge 0 \quad \text{y} \quad \Delta v(t^*, x^*) \le 0.$$

Esto implica que

$$\frac{\partial v}{\partial t}(t^*, x^*) - \Delta v(t^*, x^*) \ge 0,$$

una contradicción. Por lo tanto v alcanza su máximo en $[\delta, t]\times\{\|x\|\le M\}$ en la frontera para cualquier $\varepsilon > 0$ y por lo tanto, u también. (Un argumento similar aplica al mínimo.) Al tomar el límite conforme $\delta \to 0$, vemos que

$$\sup_{t\le T, \|x\|\le M} |u(t,x)| \le \sup_x |f(x)| < \infty$$

(pues supusimos que f es acotada) y al tomar el límite conforme $M \to \infty$, concluimos que u es acotada.

Sea $\varepsilon \in (0,t)$. Puesto que u es acotada y satisface la ecuación de calor entonces $u(s, B_{t-s})$ es una martingala en $[0, t-\varepsilon]$ y no sólo una martingala local. Por lo tanto

$$\mathbb{E}_x(u(\varepsilon, B_{t-\varepsilon})) = \mathbb{E}_x(u(t, B_0)) = u(t,x).$$

Puesto que u es continua y acotada y $u(0,x) = f(x)$, podemos utilizar el teorema de convergencia acotada para ver que

$$\mathbb{E}_x(f(B_t)) = u(t,x) . \qquad \square$$

Ejercicio 4.3. Pruebe que inversamente, si definimos

$$u(t,x) = \mathbb{E}_x(f(B_t)) = \frac{1}{(2\pi t)^{\delta/2}} \int f(y)\, f_t(y-x) ,$$

donde f_t es la densidad de B_t cuando comienza en 0 y f es medible y satisface

$$\int |f(y)|\, e^{-ay^2}\, dy < \infty$$

para toda $a > 0$, entonces u es de clase C_∞ en $(0,\infty) \times \mathbb{R}^\delta$, continua en $[0,\infty) \times \mathbb{R}^\delta$ y satisface el problema (6). *Sugerencia:* haga el cambio de variable $z = y - x$ en la definición de u y derive bajo la integral.

Generalizaremos ahora el razonamiento anterior para obtener la formulación moderna de la fórmula de Feynman-Kac. Como se observa en [**Ste01**], la fórmula de Feynman-Kac se comprende muy bien cuando se comienza con el movimiento browniano matado en un tiempo exponencial. En efecto, si T es exponencial de parámetro λ e independiente de B y definimos

$$\tilde{B}_t = \begin{cases} B_t & t < T \\ \Delta & T \geq t \end{cases}$$

(donde Δ se interpreta como el estado cementerio y extendemos a cualquier función real como cero en Δ) entonces para cualquier función continua y acotada se tiene que la función

$$u(t,x) = \mathbb{E}_x\left(f\left(\tilde{B}_t\right)\right) = e^{-\lambda t}\mathbb{E}_x(f(B_t))$$

satisface el problema de Cauchy

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{1}{2}\Delta u = \lambda u \\ u(0,x) = f(x) \end{cases}$$

En un caso más general, consideremos a

$$u(t,x) = \mathbb{E}_x\left(f(B_t)\, e^{-\int_0^t v(B_s)\, ds}\right) .$$

La interpretación es que consideramos la esperanza de un Browniano matado a tasa $v(x)$ cuando se encuentra en el estado x. Si f es continua y acotada y v es no-negativa entonces u es continua y acotada. Al utilizar la propiedad de Markov vemos que

$$\mathbb{E}_x\left(f(B_t)\, e^{-\int_0^t v(B_s)\, ds} \,\middle|\, \mathscr{F}_s\right) = e^{-\int_0^s v(B_r)\, dr} u(t-s, B_s)$$

para $s \leq t$. Definamos

$$\Pi_t = e^{-\int_0^t v(B_s)\, ds}.$$

Si u fuera de clase C_2 entonces la fórmula de Itô nos diría que

$$\Pi_t u(t-s, B_s) = u(t,x) + \int_0^s \Pi_r D_2 u(t-r, B_r)\, dB_r - \int_0^s \Pi_r D_1 u(t-r, B_r)\, dr$$
$$+ \frac{1}{2}\int_0^r \Pi_r \Delta u(t-r, B_r)\, dr - \int_0^s u(t-r, B_r)\, v(B_r)\, \Pi_r\, dr.$$

Así, vemos que una condición natural para que $\Pi_s u(t-s, B_s)$ sea una martingala local es que u satisfaga la ecuación

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{1}{2}\Delta u = vu \\ u(0,x) = f(x) \end{cases}.$$

Por otro lado, mostremos que hay a lo más una solución acotada para la ecuación anterior. En efecto, si u es una solución continua y acotada a dicha ecuación entonces la fórmula de Itô nos dice que

$$\Pi_s u(t-s, B_s)$$

es una martingala acotada. Por lo tanto

$$u(t,x) = \mathbb{E}_x(\Pi_t u(0, B_t)) = \mathbb{E}_x(\Pi_t f(B_t)).$$

Ahora pasaremos a la utilización de las ideas de Kac para la obtención de la ley arcoseno de Paul Lévy. Sea

$$A_t = \int_0^t v(B_s)\, ds$$

donde B es un movimiento browniano y sea G_t^x la función de distribución de A_t bajo la medida $\mathbb{P}_x$ que hace que el browniano comience en x . Entonces

$$\int_0^\infty \int_0^\infty e^{-\alpha t - \beta y}\, G_t(dy)\, dt = \mathbb{E}_x\left(\int_0^\infty e^{-\alpha t - \beta A_t}\, dt\right).$$

Denotaremos a la anterior esperanza por $F(x)$ (que por supuesto depende también de α y β) y sean

$$\Pi_s = e^{-\beta A_s} \quad \text{y} \quad S_t = \int_0^t e^{-\alpha s}\Pi_s\, ds,$$

por lo que

$$F(x) = \mathbb{E}_x(S_\infty).$$

Ahora entontraremos una ecuación diferencial satisfecha por F. En efecto, por la propiedad de Markov aplicada al tiempo t, vemos que

$$\mathbb{E}_x(S_\infty \mid \mathscr{F}_t) = S_t + e^{-\alpha t}\Pi_t F(B_t),$$

por lo que el proceso del lado derecho es una martingala. Supongamos que F es de clase C_2. Al aplicar la fórmula de Itô, vemos que

$$
\begin{aligned}
&dS_t + de^{-\alpha t}\Pi_t F(B_t) \\
&= e^{-\alpha t}\Pi_t \, dt + F(B_t)\left[-\beta e^{-\alpha t}\Pi_t \mathbf{1}_{B_t>0} - \alpha e^{-\alpha t}\Pi_t\right] dt \\
&+ e^{-\alpha t}\Pi_t\left[F'(B_t)\; dB_t + \frac{1}{2}F''(B_t)\; dt\right].
\end{aligned}
$$

Así, es natural buscar a F entre las soluciones a

$$\frac{1}{2}F''(x) + 1 = F(x)\left(\alpha + \beta v(x)\right).$$

Kac trata el caso particular $v(x) = x^2$. Sin embargo, podemos hacer algunas consideraciones generales: primero, existe a lo más una solución acotada F a la ecuación anterior. En efecto, si F es una solución acotada entonces

$$S_t + e^{-\alpha t}\Pi_t F(B_t)$$

es una martingala acotada y por lo tanto

$$\mathbb{E}_x(S_\infty) = F(x).$$

Por otra parte, si v es además Lipschitz continua y con valores entre 0 y 1, entonces además existe una solución acotada. Utilizaremos el siguiente argumento, puesto que

$$\alpha F \leq \frac{1}{2}F'' + 1 \leq (\alpha + \beta)\, F$$

entonces se sigue que si $F(0) > 0$ y $F'(0) > 0$ entonces

$$
\begin{aligned}
\left[F(0) - \frac{F'(0)}{\sqrt{2\alpha}}\right] + \frac{F'(0)}{\sqrt{2\alpha}}e^{\sqrt{\alpha}x} &\leq F(x) \\
&\leq \left[F(0) - \frac{F'(0)}{\sqrt{2\left(\alpha+\beta\right)}}\right] + \frac{F'(0)}{\sqrt{2\left(\alpha+\beta\right)}}e^{\sqrt{(\alpha+\beta)}x}.
\end{aligned}
$$

Así, F será acotada en $(-\infty, 0)$. Para construir soluciones acotadas en $(0, \infty)$ sólamente notamos que si $\tilde{F}(x) = F(-x)$ entonces $\tilde{F}(x)$ satisface la misma ecuación diferencial salvo que F pero

$$\tilde{F}(0) = F(0) \quad \text{y} \quad \tilde{F}'(0) = -F(0).$$

Finalmente, podemos pegar dos soluciones F y G, la primera acotada en $(-\infty, 0)$ y la segunda en $(0, \infty)$ siempre y cuando se satisfagan las relaciones

$$F(0) = G(0) > 0, F'(0) = G'(0) > 0.$$

En este caso, la solución resultante será acotada y de clase C_2 pues v es continua.

Cuando $v = \mathbf{1}_{(0,\infty)}$, las soluciones a la ecuación

$$\frac{1}{2}F''(x) + 1 = F(x)\left(\alpha + \beta \mathbf{1}_{x>0}\right)$$

no pueden ser de clase C_2, serán de clase C_2 en $\mathbb{R}\setminus\{0\}$ y en 0 tendrán derivadas laterales por la derecha y por la izquierda que difieren en $2F(0)\,\beta$. Olvidémonos por lo pronto de este detalle y notemos que la solución (acotada) a la ecuación anterior está dada por

$$F(x) = \begin{cases} Ae^{x\sqrt{2\alpha}} + \frac{1}{\alpha} & x < 0 \\ Be^{-x\sqrt{2(\alpha+\beta)}} + \frac{1}{\alpha+\beta} & x > 0 \end{cases}.$$

Si queremos que dicha función sea de clase C_1 necesitamos que se cumpla el sistema de ecuaciones

$$\begin{cases} A + \frac{1}{\alpha} = B + \frac{1}{\alpha+\beta} \\ \sqrt{2\alpha}A = -\sqrt{2\,(\alpha+\beta)}B. \end{cases}$$

Vemos que entonces

$$F(0) = \frac{1}{\sqrt{\alpha}\sqrt{\alpha+\beta}}.$$

Por lo tanto:

$$\mathbb{E}\left(\int_0^\infty e^{-\alpha t-\beta A_t}\,dt\right) = \frac{1}{\sqrt{\alpha}\sqrt{\alpha+\beta}}.$$

Por otra parte, se puede comprobar directamente que

$$\frac{1}{\sqrt{\alpha}\sqrt{\alpha+\beta}} = \int_0^\infty e^{-\alpha t}\int_0^t \frac{e^{-\beta\theta}}{\pi\sqrt{\theta(t-\theta)}},$$

por lo que se deduce que A_t/t tiene distribución Beta de parámetros $1/2$ y $1/2$. Sin embargo falta ver por qué $\mathbb{E}_x(S_\infty) = F(0)$. Esto se deduce por aproximación: sea v_n una sucesión de funciones Lipschitz continuas con valores en $[0,1]$ y que decrecen a $\mathbf{1}_{[0,\infty)}$. Entonces

$$\mathbb{E}_x\left(\int_0^\infty e^{-\alpha s-\beta\int_0^s v_n(B_r)\,dr}\right) \to \mathbb{E}_x\left(\int_0^\infty e^{-\alpha s-\beta\int_0^s \mathbf{1}_{B_r\geq 0}\,dr}\right)$$

conforme $n\to\infty$. Sea F_n la única solución acotada y no-negativa a

$$\frac{1}{2}F_n'' + 1 = F_n v_n.$$

Puesto que $0\leq F_n\leq 1/(\alpha+\beta)$ entonces F_n'' es uniformemente acotada y por lo tanto F_n' y F_n son uniformemente acotadas y equicontinuas en compactos. Si n_k es una subsucesión tal que F_{n_k}' y F_{n_k} convergen uniformemente en compactos con límite F entonces

$$\frac{1}{2}F_n''(x) + 1 \to F(x)\,\mathbf{1}_{x\geq 0}$$

para $x\neq 0$. Por lo tanto, F es de clase C_2 en $\mathbb{R}\setminus\{0\}$ y satisface

$$\frac{1}{2}F''(x) + 1 = F(x)\,\mathbf{1}_{x\geq 0}$$

si $x \neq 0$. Puesto que F es acotada y no-negativa, hemos visto que forzosamente $F(0) = \frac{1}{\sqrt{\alpha}\sqrt{\alpha+\beta}}$ y que

$$F(0) = \mathbb{E}_0\left(\int_0^\infty e^{-\alpha s-\beta\int_0^s \mathbf{1}_{B_r\geq 0}\,dr}\right).$$

8. El teorema de Girsanov

La fórmula de Itô nos dice que la clase de semimartingalas es invariante ante composición con funciones de clase C_2. Ahora examinaremos otra propiedad de invariancia de las semimartingalas: la invariancia ante cambios de medida (localmente) absolutamente continuos. Si $\mathbb{P}$ y $\mathbb{Q}$ son medidas de probabilidad absolutamente continuas y X es una semimartingala al utilizar la medida de probabilidad entonces el célebre teorema de Girsanov nos ayudará a encontrar la descomposición de semimartingala de X cuando se utiliza la medida $\mathbb{Q}$.

Sea $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad dotado de una filtración $(\mathscr{F}_t, t\geq 0)$ que satisface las condiciones habituales. Recordemos que una medida de probabilidad $\mathbb{Q}$ en $(\Omega, \mathscr{F})$ es **absolutamente continua** respecto de $\mathbb{P}$, denotado $\mathbb{Q} \ll \mathbb{P}$, si para todo $A \in \mathscr{F}$ con $\mathbb{P}(A) = 0$ se tiene que $\mathbb{Q}(A) = 0$.

Definición. Decimos que $\mathbb{Q}$ es **localmente absolutamente continua** respecto de $\mathbb{P}$, y lo denotamos por $\mathbb{Q} \lhd \mathbb{P}$, si para cada $t \geq 0$, la restricción $\mathbb{Q}|_{\mathscr{F}_t}$ de $\mathbb{Q}$ a $\mathscr{F}_t$ es absolutamente continua respecto de $\mathbb{P}|_{\mathscr{F}_t}$.

Analizaremos ahora una construcción de medidas localmente absolutamente continuas. Sea M una martingala (en $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$) no-negativa que comienza en 1. Para cada $t \geq 0$, podemos definir una medida $\mathbb{Q}_t$ en $\mathscr{F}_t$ mediante la fórmula

$$\mathbb{Q}_t(A) = \mathbb{E}(\mathbf{1}_A M_t).$$

La propiedad de martingala de M nos dice que si $A \in \mathscr{F}_s$ con $s \leq t$ entonces

$$\mathbb{Q}_t(A) = \mathbb{Q}_s(A).$$

Si M es una martingala uniformemente integrable, entonces M_t converge casi seguramente y en L_1 conforme $n \to \infty$ a una variable integrable M_∞ y se tiene que

$$\mathbb{Q}_t(A) = \mathbb{E}(\mathbf{1}_A M_\infty),$$

por lo que en este caso, existe una medida $\mathbb{Q}$ tal que

$$\mathbb{Q}|_{\mathscr{F}_t} = \mathbb{Q}_t.$$

Sin embargo, esto es válido aún cuando M no sea uniformemente integrable. En efecto, por la propiedad de martingala podemos definir una función $\mathbb{Q} : \cup_t \mathscr{F}_t \to [0, \infty)$ tal que

$$\mathbb{Q}|_{\mathscr{F}_t} = \mathbb{Q}_t.$$

A veces, $\mathbb{Q}$ se puede extender a una medida $\mathbb{Q}$ en $\mathscr{F}_\infty = \sigma(\mathscr{F}_t : t \geq 0)$ (por ejemplo, si nuestro espacio de probabilidad es un espacio de Borel estándar, como en [**Par05**]). Por esta razón, asumiremos en esta sección que $\mathscr{F} = \mathscr{F}_\infty$ y es claro que entonces $\mathbb{Q} \lhd \mathbb{P}$.

De hecho, la construcción anterior se puede modificar en caso de que M sea una martingala local continua no-negativa. Si (T_n) es una sucesión de tiempos de paro que localice a M, se puede definir

$$\mathbb{Q}^n(A) = \mathbb{E}(\mathbf{1}_A M_{T_n}) \quad \text{para} \quad A \in \mathscr{F}_{T_n}.$$

Puesto que $T_n \to \infty$, se sigue que $\sigma(\mathscr{F}_{T_n} : n \geq 1) = \mathscr{F}_\infty$. De la misma manera, en ciertas circunstancias es posible encontrar una medida $\mathbb{Q}$ en $\mathscr{F}_\infty$ tal que $\mathbb{Q}|_{\mathscr{F}_{T_n}} = \mathbb{Q}_n$.

Proposición 4.5. *Supongamos que $\mathbb{Q} \lhd \mathbb{P}$ y sea*

$$\tilde{D}_t = \frac{d\mathbb{P}|\mathscr{F}_t}{d\mathbb{Q}|_{\mathscr{F}_t}}.$$

Entonces $\tilde{D}$ admite una modificación D que es una una martingala càd no-negativa y uniformemente integrable. Para todo T tiempo de paro se tiene:

$$D_T = \frac{d\mathbb{P}|\mathscr{F}_T}{d\mathbb{Q}|_{\mathscr{F}_T}}.$$

Si $\mathbb{Q}$ es equivalente a $\mathbb{P}$ entonces $D_t > 0$ para toda $t \geq 0$ casi seguramente.

Demostración. Si $A \in \mathscr{F}_s$ y $s \leq t$ entonces $A \in \mathscr{F}_t$ y por definición de $\tilde{D}_s$ y $\tilde{D}_t$: $\mathbb{E}\left(\mathbf{1}_A \tilde{D}_s\right) = \mathbb{Q}(A) = \mathbb{E}\left(\mathbf{1}_A \tilde{D}_t\right)$. Por lo tanto $\tilde{D}$ es una $\mathbb{P}$-martingala. Puesto que hemos asumido las condiciones habituales para $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ vemos que $\tilde{D}$ admite una modificación càdlàg que también es una martingala y también se satisface la relación

$$D_t = \frac{d\mathbb{P}|\mathscr{F}_t}{d\mathbb{Q}|_{\mathscr{F}_t}}.$$

Notemos que lo anterior vale también para $t = \infty$, por lo que D es uniformemente integrable. Si T es un tiempo de paro y $A \in \mathscr{F}_T$ entonces, al aplicar muestreo opcional, vemos que

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}(\mathbf{1}_A D_\infty) = \mathbb{E}_{\mathbb{P}}(\mathbf{1}_A D_T).$$

Por lo tanto

$$D_T = \frac{d\mathbb{P}|\mathscr{F}_T}{d\mathbb{Q}|_{\mathscr{F}_T}}.$$

Finalmente, si $S = \inf\{t \geq 0 : D_t = 0\}$ entonces

$$\mathbb{Q}(S < \infty) = \mathbb{E}_{\mathbb{P}}(\mathbf{1}_{S<\infty} D_S) = 0$$

y si $\mathbb{Q}$ es equivalente a $\mathbb{P}$, esto implica que $\mathbb{P}(S < \infty) = 0$. $\square$

Así, en el caso en que tengamos dos medidas de probabilidad equivalente, el proceso de derivadas de Radon-Nikodym es una martingala estrictamente positiva. El siguiente resultado nos permitirá expresar a dicha martingala, cuando tenga trayectorias continuas, como una exponencial estocástica.

PROPOSICIÓN 4.6. *Sea D una martingala local continua estrictamente positiva. Existe entonces una única martingala local continua L tal que $D = \mathscr{E}(L)$. Además:*

$$L_t = \log(D_0) + \int_0^t D_s^{-1}\, dD_s.$$

DEMOSTRACIÓN. Para probar la unicidad, supongamos que $D = \mathscr{E}(L) = \mathscr{E}\left(\tilde{L}\right)$. Entonces

$$1 = D\frac{1}{D} = \mathscr{E}(L)\,\frac{1}{\mathscr{E}\left(\tilde{L}\right)} = e^{L-\tilde{L}}.$$

Para la existencia, utilizamos la fórmula de Itô con la función log, que es infinitamente diferenciable en $(0,\infty)$ a la martingala local continua estríctamente positiva D. Se tiene entonces que

$$\log(D_t) = \log(D_0) + \int_0^t D_s^{-1}\, dD_s - \frac{1}{2}\int_0^t D_s^{-2}\, d\langle D\rangle_s.$$

Si

$$L_t = \log(D_0) + \int_0^t D_s^{-1}\, dD_s,$$

entonces

$$\langle L\rangle_t = \int_0^t D_s^{-2}\, d\langle D\rangle_s,$$

por lo que

$$\log(D_t) = L_t - \frac{1}{2}\langle L\rangle_t$$

y por lo tanto

$$D_t = \exp L_t - \frac{1}{2}\langle L\rangle_t = \mathscr{E}(L)_t\,. \qquad \square$$

Ahora podemos enunciar el teorema de Girsanov.

TEOREMA 4.11 (Teorema de Girsanov). *Sea $\mathbb{Q}$ equivalente a $\mathbb{P}$ en $\mathscr{F}_\infty$. Sea D la versión càdlàg del proceso de derivadas de Radon-Nikodym y supongamos que D es continuo. Sea L una martingala local continua tal que $D = \mathscr{E}(L)$. Si M es cualquier $(\mathscr{F}_t,\mathbb{P})$-martingala local continua el proceso $\tilde{M}$ dado por*

$$\tilde{M}_t = M_t - \langle M, L\rangle_t$$

es una $(\mathscr{F}_t,\mathbb{Q})$-martingala local continua.

Notemos que en particular, M es una $\mathbb{Q}$-semimartingala. Notemos además que la variación cuadrática no depende de la medida que estemos utilizando puesto que los límites en probabilidad coinciden para medidas equivalentes. Así, si M es un movimiento browniano bajo $\mathbb{P}$, entonces $\tilde{M}$ lo es bajo $\mathbb{Q}$.

Demostración. Mostremos primero que si XD es una $\mathbb{P}$-martingala local continua entonces X es una $\mathbb{Q}$-martingala local. En efecto,

$$T_n = \inf\{t \geq 0 : D_t \geq n \text{ ó } |X_t| \geq n\}.$$

Entonces (T_n) es una sucesión creciente de tiempos de paro que convergen a ∞, $(XD)^{T_n}$ es una martingala acotada y X^{T_n} es un proceso acotado. Por lo tanto, si $A \in \mathscr{F}_s$ y $s \leq t$ entonces

$$\mathbb{E}_{\mathbb{Q}}(XT_n \wedge t\mathbf{1}_A) = \mathbb{E}_{\mathbb{P}}(X_{t\wedge T_n} D_{t\wedge T_n} \mathbf{1}_A) = \mathbb{E}_{\mathbb{P}}(X_{s\wedge T_n} D_{s\wedge T_n} \mathbf{1}_A) = \mathbb{E}_{\mathbb{Q}}(X_{T_n\wedge s} \mathbf{1}_A).$$

Por otra parte, notemos que puesto que $\mathbb{P}(T_n \leq t) \to 0$ se sigue que $\mathbb{Q}(T_n \leq t) \to 0$ y que por lo tanto $T_n \to \infty$ $\mathbb{Q}$-casi seguramente. Así, X es una $\mathbb{Q}$ martingala local continua.

Ahora aplicaremos la observación anterior. Si M es una $\mathbb{P}$-martingala local continua y $\tilde{M} = M - \langle M, L\rangle$, podemos aplicar la fórmula de Itô para escribir

$$\begin{aligned} D_t\tilde{M} &= D_0\tilde{M}_0 + \int_0^t D_s\,d\tilde{M}_s + \int_0^t \tilde{M}_s\,dD_s + \langle \tilde{M}, D\rangle \\ &= D_0\tilde{M}_0 + \int_0^t D_s\,dM_s + \int_0^t \tilde{M}_s\,dD_s - \int_0^t D_s\,d\langle M, L\rangle_s + \langle M, D\rangle \\ &= D_0\tilde{M}_0 + \int_0^t D_s\,dM_s + \int_0^t \tilde{M}_s\,dD_s. \end{aligned}$$

Por lo tanto $D\tilde{M}$ es una $\mathbb{P}$-martingala local continua y se deduce que entonces $\tilde{M}$ es una $\mathbb{Q}$-martingala local continua. □

Uno de los ejemplos típicos de aplicación del teorema de Girsanov es al movimiento browniano con deriva. En efecto, si B es un movimiento browniano bajo $\mathbb{P}$ y

$$\mathbb{Q}_t(A) = \mathbb{E}_{\mathbb{P}}\left(\mathbf{1}_A e^{\mu B_t - \mu^2 t/2}\right),$$

para $A \in \mathscr{F}$, entonces $\mathbb{Q}_t$ es una medida de probabilidad absolutamente continua respecto de $\mathbb{P}$ y equivalentemente a ella en $\mathscr{F}_t$. Por lo tanto, bajo $\mathbb{Q}_t$ el proceso $B-\mu\,\mathrm{Id}$ es un browniano en $[0, t]$; equivalentemente, bajo $\mathbb{Q}_t$, B es un browniano con deriva $-\mu$ en $[0,t]$. Otro ejemplo es que si $T_b(X) = \inf\{t \geq 0 : X_s \geq b\}$ entonces

$$\begin{aligned}\mathbb{E}\left(e^{-\lambda T_b(B+\mu \operatorname{Id})}\right) &= \lim_{t\to\infty} \mathbb{E}\left(e^{-\lambda T_b(B+\mu \operatorname{Id})\wedge t}\right) \\ &= \lim_{t\to\infty} \mathbb{E}\left(e^{-\lambda T_b(B)\wedge t} e^{\mu B_{T_b(B)\wedge t} - \mu^2 T_b(B)\wedge t/2}\right) \\ &= \mathbb{E}\left(e^{-\lambda T_b(B)} e^{\mu b - \mu^2 T_b(B)/2}\right) \\ &= e^{\mu b} e^{-|b|\sqrt{2\lambda+\mu^2}}.\end{aligned}$$

En particular, vemos que

$$\mathbb{P}(T_b(B+\mu \operatorname{Id}) < \infty) = \lim_{\lambda\to 0} \mathbb{E}\left(e^{-\lambda T_b(B+\mu \operatorname{Id})}\right) = e^{\mu b - \mu|b|} = \begin{cases} = 1 & \operatorname{sgn}(b) = \operatorname{sgn}\mu \\ < 1 & \text{otro caso} \end{cases}.$$

Una extensión de la idea anterior permite resolver el siguiente problema.

Ejercicio 4.4. Sea B un movimiento browniano que comienza en cero y $\gamma \in \mathbb{R}$. Sea

$$T = \inf\{t \geq 0 : |B_t + \gamma t| = 1\}.$$

(1) Pruebe que si $\gamma = 0$ entonces T y B_T son independientes.
(2) Al utilizar el teorema de Girsanov muestre la independencia entre T y B_T cuando $\gamma \neq 0$.

Otro ejemplo de aplicación es el siguiente. Notemos que

$$\mathbb{E}\left(f\left(\sup_{s\leq t} B_s + \mu_s\right)\right) = \mathbb{E}\left(f\left(\sup_{s\leq t} B_s\right) e^{-\mu B_t}\right).$$

En particular

$$\mathbb{P}\left(\sup_{s\leq t} B_s + \mu s \in dx, B_t \in dy\right) = \mathbb{P}\left(\sup_{s\leq t} B_s \in dx, B_t \in dy\right) e^{\mu y - \mu^2 t/2}.$$

Este es un resultado no trivial puesto que se conoce explícitamente la densidad conjunta de $\left(B_t, \sup_{s\leq t} B_s\right)$:

$$\mathbb{P}\left(\sup_{s\leq t} B_s \in dy, B_t \in dx\right) = \frac{2\,(2y-x)}{\sqrt{2\pi t^3}} e^{-(2y-x)^2/2t} \mathbf{1}_{y>0, x\leq y}.$$

Una de las aplicaciones del teorema de Girsanov es a la técnica de remoción de deriva.

Ejercicio 4.5. Considere la ecuación diferencial estocástica

$$dX_t = dB_t + b(X_t)\, dt \quad X_0 = x \tag{7}$$

donde b es medible y acotada. Suponga que bajo P, X es un movimiento browniano que comienza en x. Utilice el teorema de Girsanov para encontrar una medida de probabilidad $\tilde{\mathbb{P}}$ tal que si definimos a

$$B_t = X_t - \int_0^t b(X_s)\ ds$$

entonces $(B_t)_{t\leq 1}$ sea un movimiento browniano bajo $\tilde{\mathbb{P}}$. Note que X resuelve entonces la ecuación diferencial estocástica (7); esta solución es llamada solución por transformación de deriva.

CAPÍTULO 5

Integral estocástica respecto de semimartingalas

1. Integrales de Lebesgue-Stieltjes respecto de procesos de Poisson

Sea $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad y $(\mathscr{F}_t, t \geq 0)$ una filtración estándar. Al considerar a los procesos estocásticos como definidos en $[0, \infty) \times \mathbb{R}$, pensamos a los procesos como funciones aleatorias.

Definición. La **σ-álgebra previsible**, denotada $\mathscr{P}$, es la generada en $\mathbb{R}_+ \times \Omega$ por los procesos estocásticos continuos por la izquierda y adaptados a $(\mathscr{F}_t)$.

Proposición 5.1. *La σ-álgebra previsible coincide con la generada por:*

(1) *Los conjuntos de la forma $A \times (s, t]$ donde $0 \leq s < t < \infty$ y $A \in \mathscr{F}_s$.*
(2) *Los procesos adaptados elementales.*
(3) *Los procesos continuos y adaptados.*

Pasaremos ahora a los procesos de Poisson y de Poisson compuesto.

Definición. Un **$(\mathscr{F}_t)$-proceso de Poisson** es un proceso càd, adaptado, que comienza en cero y tal que para $s < t$ y $k \in \mathbb{N}$ se tiene que

$$\mathbb{P}(N_t - N_s = k \mid \mathscr{F}_s) = e^{-\lambda(t-s)} \frac{[\lambda(t-s)]^k}{k!}.$$

En otras palabras, la distribución condicional del incremento $N_t - N_s$ dada $\mathscr{F}_s$ es Poisson de parámetro $\lambda(t-s)$.

Proposición 5.2. *Un proceso estocástico N es un $(\mathscr{F}_t)$-proceso de Poisson si y sólo si es un $(\mathscr{F}_t)$-proceso de Lévy con trayectorias constantes por pedazos y con saltos de magnitud* 1.

Ahora veremos como producir martingalas a partir de integrales respecto del proceso de Poisson. Sea N un $(\mathscr{F}_t)$-proceso de Poisson de intensidad λ. Al proceso M dado por $M_t = N_t - \lambda t$ lo llamamos proceso de Poisson compensado asociado a N.

Proposición 5.3. *Los procesos M y $M_t^2 - \lambda t$ son $(\mathscr{F}_t)$-martingalas. Si Z es un proceso previsible tal que*

$$\mathbb{E}\left(\int_0^t |Z_s| \, ds\right) < \infty$$

para toda $t > 0$, entonces el proceso $Z \cdot M$ dado por

$$Z \cdot M_t = \int_0^t Z_s \, dM_s$$

es una $(\mathscr{F}_t)$-martingala.

Demostración. Las primeras dos afirmaciones son estándar y valen más generalmente para procesos de Lévy centrados.

Para la tercera afirmación, sea Z un proceso previsible elemental. Es decir, Z admite la descomposición

$$Z_t(\omega) = Z_0(\omega) \, \mathbf{1}_{t=0} + \sum_{i=1}^{n} H_i(\omega) \, \mathbf{1}_{t \in (t_{i-1}, t_i]}$$

donde $0 = t_0 < t_1 < \cdots < t_n$ y H_i es $\mathscr{F}_{t_{i-1}}$-medible y acotada. Entonces

$$(Z \cdot M)_t = \sum_{i=1}^{n} H_i \left(M_{t_i \wedge t} - M_{t_{i-1} \wedge t} \right).$$

Puesto que M es martingala y H_i es $\mathscr{F}_{t_{i-1}}$-medible, vemos que $Z \cdot M$ es una martingala. Si Z es un proceso previsible que satisface la condición de integrabilidad enunciada. □

Notemos que la afirmación anterior en realidad es válida si M es una martingala de variación finita

2. Medidas de Poisson aleatorias

Definición. Sea $(S, \mathscr{S})$ un espacio medible. Una medida aleatoria es una aplicación $\Xi : \Omega \times \mathscr{S} \to [0, \infty]$ tal que

(1) para toda $\omega \in \Omega$ la aplicación $A \mapsto \Xi(\omega, A)$ es una medida en $\mathscr{S}$ y
(2) para toda $A \in \mathscr{S}$ la aplicación $\omega \mapsto \Xi(\omega, A)$ es una variable aleatoria.

Definición. Sea $(S, \mathscr{S})$ un espacio medible y ν una medida. Una medida aleatoria Ξ es una **medida de Poisson aleatoria** de medida de intensidad ν si

(1) para todo $A \in \mathscr{S}$ la variable $\Xi(A)$ tiene distribución Poisson de parámetro $\nu(A)$ y
(2) si $A_1, \ldots, A_n \in \mathscr{S}$ son ajenos por pares entonces $\Xi(A_1), \ldots, \Xi(A_n)$ son independientes.

Teorema 5.1. *Sea ν una medida σ-finita en el espacio medible $(S, \mathscr{S})$. Entonces existe (un espacio de probabilidad en el que está definida) la medida de Poisson aleatoria con medida de intensidad ν.*

Proposición 5.4. *Sea ν una medida σ-finita en $(S, \mathscr{S})$ una medida de Poisson aleatoria con medida de intensidad ν.*

(1) *(Fórmula exponencial) Si $f : S \to \mathbb{R}_+$ es medible entonces la integral de f respecto de Ξ, denotada por Ξf, es una variable aleatoria y*

$$\mathbb{E}\left(e^{-\Xi f}\right) = e^{-\int \left(1-e^{-f}\right) d\nu}.$$

(2) *La variable aleatoria Ξf es casi seguramente finita o casi seguramente infinita de acuerdo a si la integral $\int 1 \wedge f \, d\nu$ es finita o no.*

Consideremos un parámetro de intensidad $\lambda > 0$ y una distribución de salto μ en $\mathbb{R}$. Sea N un proceso de Poisson de parámetro $\lambda > 0$ independientemente de la sucesión iid $X_1, X_2, \ldots$ donde X_i tiene distribución μ. Sabemos que el proceso de Poisson compuesto se construye como

$$X_t = \sum_{i \leq N_t} X_i.$$

Una interpretación en términos de medidas aleatorias es la siguiente: sean $T_1 < T_2 < \cdots$ los instantes de salto del proceso N y consideremos a la medida

$$\Xi = \sum_n \delta_{(T_n, X_n)}.$$

Puesto que $\sum_n \delta_{T_n}$ es una medida de Poisson aleatoria de intensidad λLeb es fácil ver que Ξ es una medida de Poisson aleatoria de intensidad $\lambda\mathrm{Leb} \times \mu$. Además, si $f_t : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ está dada por $f_t(s, x) = \mathbf{1}_{s \leq t} x$ entonces

$$X_t = \Xi f_t.$$

Elaboraremos esta idea para construir a una clase importante de procesos de Levy.

Definición. Un **subordinador** es un proceso de Lévy X tal que casi seguramente las trayectorias $t \mapsto X_t$ son no-decrecientes.

Ejercicio 5.1. Pruebe que un proceso de Lévy X es un subordinador si y sólo si la distribución de X_t tiene soporte en $[0, \infty)$ para alguna $t > 0$. *Sugerencia:* comience el ejercicio en el caso en que el soporte esté contenido en $[0, \infty)$ para toda $t > 0$.

Ahora daremos una construcción de un subordinador mediante una medida aleatoria de Poisson. Sean μ una medida en $(0, \infty)$ tal que

$$\int_0^\infty 1 \wedge x \, \mu(dx) < \infty.$$

y $d \geq 0$. Sea Ξ una medida de Poisson aleatoria en $\mathbb{R}_+ \times \mathbb{R}_+$ con intensidad $\mathrm{Leb} \times \mu$. Para las funciones f_t que definimos anteriormente, definamos a

$$X_t = dt + \Xi f_t = dt + \int_0^t \int_0^\infty x \, \Xi(ds, dx).$$

Proposición 5.5. *El proceso X es un subordinador. Si μ es una medida infinita entonces el conjunto de discontinuidades de X es denso y numerable. Además, las distribuciones unidimensionales de X están caracterizadas por*

$$\mathbb{E}\left(e^{-\lambda X_t}\right) = e^{-t\Phi(\lambda)} \quad \textit{donde} \quad \Phi(\lambda) = d + \int_0^\infty \left(1 - e^{-\lambda x}\right) \mu(dx)\,.$$

Demostración. Primero notemos que puesto que $s \leq t$ implica $f_s \leq f_t$ entonces X tiene trayectorias no-decrecientes. Además, nuestra hipótesis sobre μ implica que $X_n < \infty$ para toda $n \in \mathbb{N}$ casi seguramente y por lo tanto $X_t < \infty$ para toda $t \geq 0$ casi seguramente. Claramente $X_0 = 0$. Si $t \downarrow s$ entonces $f_t \downarrow f_s$ y por lo tanto X tiene trayectorias continuas por la derecha. Puesto que X es no decreciente, tiene límites por la izquierda. Si (t_n, x_n) es una enumeración de los átomos de Ξ se sigue que X tiene una discontinuidad en t_n y que $X_{t_n} - X_{t_n-} = x_n$. Si μ es una medida infinita entonces $\Xi([0,t] \times \mathbb{R}_+) = \infty$ casi seguramente por lo que X tiene una cantidad numerable de discontinuidades en cada intervalo compacto. Esto implica que las discontinuidades de X son densas.

Sean $0 = t_0 < t_1 < \cdots < t_n$ y $\lambda_1, \ldots, \lambda_n \geq 0$. Al aplicar la fórmula exponencial a la función

$$f(s,x) = \sum_i \mathbf{1}_{t_i \leq s \leq t} \lambda_i x$$

se tiene que

$$\Xi f = \sum_{i=1}^n \lambda_i \left(X_{t_i} - X_{t_{i-1}}\right)$$

y por lo tanto:

$$\mathbb{E}\left(e^{-\sum_{i=1}^n \lambda_i \left(X_{t_i} - X_{t_{i-1}}\right)}\right) = e^{-\sum_{i=1}^n (t_i - t_{i-1}) \int \left(1 - e^{-\lambda_i x}\right) \mu(dx)}$$

Al utilizar $\lambda_j = 0$ si $j \neq i$ nos damos cuenta de que

$$\mathbb{E}\left(e^{-\lambda_i \left(X_{t_i} - X_{t_{i-1}}\right)}\right) = e^{-(t_i - t_{i-1}) \int \left(1 - e^{-\lambda_i x}\right) \mu(dx)}$$

y que por lo tanto los incrementos de X son independientes y estacionarios. Claramente X comienza en cero, por lo que resta probar que X tiene trayectorias càdlàg. □

Corolario 10. *Se tiene que $\mathbb{E}(X_t) = t \int x\, \mu(dx)$. Si dicha cantidad es finita, $Var(X_t) = t \int_0^\infty x^2\, \mu(dx)$.*

Ahora veremos como el fenómeno de compensación de medidas aleatorias nos ayuda a extender la construcción de subordinadores a procesos de Lévy con trayectorias que no sean monótonas. Sea ν una medida en $[-1,1] \setminus \{0\}$ tal que

$$\int x^2\, \nu(dx) < \infty.$$

Sea Ξ una medida de Poisson aleatoria en $[0,\infty)\times[-1,1]$ de intensidad $\lambda\times\nu$. Consideremos, para $\varepsilon>0$ al conjunto $I_\varepsilon=[-1,1]\setminus[-a,a]$. Construyamos al proceso

$$X_t^\varepsilon=\int x\mathbf{1}_{|x|\in I_\varepsilon}\mathbf{1}_{s\leq t}\,\Xi(ds,dx)-t\nu([0,t]\times I_\varepsilon)\,.$$

Puesto que ν es inita en I_ε, X^ε no es más que un proceso de Poisson compuesto compensado. Esto implica que X^ε es una martingala. Además, vemos que si $\varepsilon'<\varepsilon$

$$\mathrm{Var}\Big(X_t^\varepsilon-X_t^{\varepsilon'}\Big)=t\int_{I_{\varepsilon'}\setminus I_\varepsilon}x^2\,\nu(dx)\,.$$

Por la hipótesis sobre ν se sigue que

$$\lim_{\varepsilon\to0}\sup_{\varepsilon'\leq\varepsilon}\mathrm{Var}\Big(X_t^\varepsilon-X_t^{\varepsilon'}\Big)=0.$$

Al utilizar la desigualdad de Doob vemos que

$$\lim_{\varepsilon\to0}\sup_{\varepsilon'\leq\varepsilon}\mathbb{E}\bigg(\sup_{s\leq t}\Big|X_t^\varepsilon-X_t^{\varepsilon'}\Big|^2\bigg)=0.$$

Se concluye que $X^\varepsilon,\varepsilon>0$ es una sucesión de Cauchy uniforme en L_2 y por lo tanto existe una sucesión $\varepsilon_n\downarrow0$ tal que X^{ε_n} converge uniformemente en compactos a un proceso càd X. X debe ser un proceso de Lévy y al utilizar el teorema de convergencia dominada vemos que

$$\mathbb{E}\big(e^{uX_t}\big)=e^{t\Psi(\lambda)}\quad\text{donde}\quad\Psi(\lambda)=\int_{[-1,1]}(e^{ux}-1-x)\ \nu(dx)\,.$$

3. Procesos de Lévy y la descomposición de Lévy-Itô

4. La descomposición de Doob-Meyer

En esta sección mostraremos el resultado en el que se puede basar una teoría de integración estocástica para martingalas càdlàg: la descomposición de Doob-Meyer.

Definición. Un proceso estocástico adaptado $X=(X_t,t\leq1)$ es de clase D si la familia

$$\{X_\tau:\tau\text{ es tiempo de paro acotado por }1\}$$

es uniformemente integrable.

Teorema 5.2. *Sea S una submartingala càdlàg en $[0,1]$ de clase D. Entonces existe una martingala M y un proceso predecible no decreciente A que comienza en cero tal que*

$$S=M+A.$$

El teorema anterior admite una versión en $[0, \infty)$ si se asume que S es de clase DL: para cualquier $T > 0$, la familia

$$\{S_\tau : \tau \text{ es tiempo de paro acotado por } T\}.$$

El teorema anterior admite también una versión muy sencilla a tiempo discreto conocida como la descomposición de Doob de una martingala. Utilizaremos esta versión junto con un procedimiento para pasar al limite para verificar la existencia. La unicidad será consecuencia del siguiente resultado.

Proposición 5.6. *Si M es una martingala predecible de variación acotada entonces tiene trayectorias constantes.*

Demostración. Definamos a la medida con signo μ en la σ-álgebra predecible mediante la fórmula

$$\mu(A) = \mathbb{E}\left(\int_0^\infty \mathbf{1}_{(t,\omega)\in A}\, dM_t(\omega)\right).$$

Como M es martingala, se sigue que $\mu((t, \infty) \times B) = 0$ para cualquier $B \in \mathscr{F}_t$. Por clases monótonas, se verifica entonces que $\mu = 0$ en $\mathscr{P}$.

Puesto que M es predecible, su proceso de saltos ΔM también lo es. Así, $\mu\{t \geq 0 : \Delta M_t > 0\} = 0$ y $\mu\{t \geq 0 : \Delta M_t < 0\} = 0$. Por lo tanto, M es continua. Puesto que M es de variación acotada, concluimos que tiene trayectorias constantes. □

Sea S una submartingala càdlàgen $[0, 1]$. Para cada $n \geq 1$, sea $\mathscr{P}_n$ el conjunto de los racionales diádicos de orden n en $[0, 1]$. Recordemos la descomposición de Doob de S sobre $\mathscr{P}_n$: definamos a $A_0 = 0$ y a

$$A_{\frac{k+1}{2^n}} = A_{\frac{k}{2^n}} + \mathbb{E}\left(S_{\frac{k+1}{2^n}} - S_{\frac{k}{2^n}} \,\middle|\, \mathscr{F}_{\frac{k}{2^n}}\right).$$

Entonces A^n es no-decreciente puesto que S es submartingala. Es claramente predecible y comienza en cero por definición. Por otra parte, $S - A^n$ es una martingala sobre $\mathscr{P}_n$ ya que

$$\mathbb{E}\left(S_{\frac{k+1}{2^n}} - S_{\frac{k}{2^n}} - A^n_{\frac{k+1}{2^n}} + A^n_{\frac{k}{2^n}} \,\middle|\, \mathscr{F}_{\frac{k}{2^n}}\right) = 0$$

por definición de A^n. Así, vemos que $S = M^n + A^n$ sobre $\mathscr{P}_n$. El detalle delicado es pasar al límite conforme $n \to \infty$. Esto se logrará mediante un resultado de espacios de Hilbert llamado Lema de Komlos.

Proposición 5.7. *Sea (f_n) una sucesión acotada en un espacio de Hilbert. Entonces existe g_n en la envolvente convexa de $\{f_n, f_{n+1}, \ldots\}$ tal que g_n converge en H.*

Demostración. Sea $A = \sup_{n\geq 1} \inf \{\|g\| : g \in \operatorname{conv}(f_n, f_{n+1}, \ldots)\}$. Entonces $A < \infty$ al ser la sucesión (f_n) acotada. Por lo tanto, para cada n se puede escoger $g_n \in \operatorname{conv}(f_n, f_{n+1}, \ldots)$ tal que $\|g_n\| \leq A + 1/n$. Si n, m son suficientemente grandes, entonces $\|(g_n+g_m)\| > A-\varepsilon$. Por la identidad del paralelogramo se tiene que

$$\|g_n - g_m\|^2 \leq 2\|g_n\|^2 + 2\|g_m\|^2 - \|g_k + g_m\|^2 \leq 4\left(A+1/n\right)^2 - 4\left(A-\varepsilon\right)^2.$$

Al ser la sucesión (g_n) de Cauchy, converge en el espacio de Hilbert. □

A partir del teorema anterior podemos obtener límites casi seguros de sucesiones uniformemente integrables módulo la consideración de envolventes convexas.

Proposición 5.8. *Sea $X_1, X_2, \ldots$ variables aleatorias uniformemente integrables. Entonces existen variables aleatorias $Y_i \in conv(X_i, X_{i+1}, \ldots)$ tales que (Y_i) converge en L_1.*

Demostración. Primero notamos que si $\tilde{X}_1, \tilde{X}_2, \ldots$ converge a $\tilde{X}$ en L_1 y $Y_n \in \operatorname{conv}\left(\tilde{X}_n, \tilde{X}_{n+1}, \ldots\right)$ entonces Y_n converge a $\tilde{X}$ en L_1. En efecto, si $Y_n = \sum_{k=1}^{N_n} \lambda_k^n \tilde{X}_k$ donde $N_n \geq n$, $\lambda_k^n \geq 0$ y $\sum_{k=1}^{N_n} \lambda_k^n = 1$ entonces

$$\mathbb{E}\left(\left|Y_n - \tilde{X}\right|\right) \leq \sum_{k=1}^{N_n} \lambda^n \mathbb{E}\left(\left|\tilde{X}_k - \tilde{X}\right|\right) \leq \sup_{k\geq n} \mathbb{E}\left(\left|\tilde{X}_k - \tilde{X}\right|\right) \to 0.$$

Para cada i, n definamos a $X_n^i = X_n \mathbf{1}_{|X_n|\leq i}$. Puesto que $\left(X_n^1, n\right)$ está acotada en L_2, existen una combinación convexa $\lambda_n^{1,n}, \ldots \lambda_{N_n^1}^{1,n}$ (con $N_n^1 \geq n$) tal que

$$\sum_{k=n}^{N_n^1} \lambda_k^{1,n} X_k^1$$

converge en L_2 y en L_1.

Ahora aplicamos la Proposición 5.7 a la sucesión

$$\left(\sum_{k=n}^{N_n^1} \lambda_k^{1,n} X_k^2, n \geq 1\right)$$

para obtener combinaciones convexas $\tilde{\lambda}_n^{2,n}, \ldots, \tilde{\lambda}_{N_n^2}^{2,n}$ (con $N_n^2 \geq n$) tales que

$$\sum_{k=n}^{N_n^2} \tilde{\lambda}_k^{2,n} \sum_{j=k}^{N_k^1} \lambda_j^{1,k} X_j^2 = \sum_{j=n} X_j^2 \sum_{k} \tilde{\lambda}_k^{2,n} \lambda_j^{1,k}$$

converge en L_2 y en L_1. Si definimos

$$\lambda_j^{2,n} = \sum_k \tilde{\lambda}_k^{2,n} \lambda_j^{1,k},$$

y a N_n^2 como el máximo k tal que $\lambda_j^{2,n} \neq 0$ entonces $\lambda_n^{2,n}, \ldots, \lambda_{N_n^2}^{2,n}$ es una combinación convexa tal que $\sum_n \lambda_k^{2,n} X_k^i$ converge para $i = 1, 2$. Obviamente, este procedimiento se puede generalizar para construir combinaciones convexas $\lambda_n^{j,n}, \ldots, \lambda_{N_n^j}^{j,n}$ para las cuales $\sum_k \lambda_k^{2,n} X_k^i$ converge en L_1 para $1 \leq i \leq j$.

Se afirma ahora que $\sum_{k=n}^{N_n^n} \lambda_k^{n,n} X_k^i$ converge para toda i. En efecto, esto sucede pues

$$\sum_{k=n}^{N_n^n} \lambda_k^{n,n} X_k^i \in \operatorname{conv}\left(\sum_{k=m}^{N_n^i} \lambda_k^{i,n} X_k^i : i \geq n \right)$$

y el lado derecho es convergente en L_1.

Finalmente, veamos que $\sum_{k=n}^{N_n^n} \lambda_k^{n,n} X_k, n \geq 1$ converge en L_1. En efecto, notemos que por integrabilidad uniforme se tiene que

$$\lim_{i\to\infty} \sup_n \mathbb{E}\big(\big| X_n^i - X_n \big| \big) = 0.$$

Por lo tanto:

$$\lim_{i\to\infty} \sup_n \mathbb{E}\left(\left| \sum_{k=n}^{N_n^n} X_k^i - X_k \right| \right) = 0.$$

Por esta razón, la sucesión $\sum_{k=n}^{N_n^n} \lambda_k^{n,n} X_k, n \geq 1$ es de Cauchy. □

Aplicaremos la Proposición 5.8 a la sucesión de martingalas M^n. Antes que nada, puesto que la filtración es estándar, podemos escoger una extensión càdlàg de M^n al poner $M_t^n = \mathbb{E}(M_1^n \mid \mathscr{F}_t)$. Para esta versión probamos el siguiente resultado.

Lema 4. *Las variables aleatorias $M_1^n, n \geq 1$ son uniformemente integrables.*

Demostración. Puesto que $S_1 = A_1^n + M_1^n$, es equivalente mostrar que la familia A_1^n es uniformemente integrable. Para verificar esto, definimos

$$\tau_r^n = \min\left\{ k/2^n : A_{k/2^n}^n > r \right\} \wedge 1.$$

Entonces, puesto que las trayectorias de A^n son no-decrecientes y por muestreo opcional obtenemos

$$\begin{aligned}
\frac{1}{2}\mathbb{E}\big(A_1^n \mathbf{1}_{A_1^n > 2r}\big) &\leq \mathbb{E}\big([A_1^n - A_1^n \wedge r]\, \mathbf{1}_{A_1^n > 2r}\big) \\
&\leq \mathbb{E}(A_1^n - A_1^n \wedge r) \\
&\leq \mathbb{E}\Big(A_1^n - A_{\tau_r^n}^n \Big) \\
&= \mathbb{E}\big(S_1 - S_{\tau_r^n}\big) \\
&= \mathbb{E}\big(\big[S_1 - S_{\tau_r^n}\big]\, \mathbf{1}_{\tau_r^n < 1}\big).
\end{aligned}$$

Al aplicar la propiedad de submartingala y la definición de A^n se obtiene

$$r\mathbb{P}(A_1^n > r) \leq \mathbb{E}(A_1^n) = \mathbb{E}(S_1)\,,$$

por lo que

$$\lim_{r\to\infty} \sup_n \mathbb{P}(A_1^n > r) = 0.$$

Puesto que las variables aleatorias $S_1^n - S_{\tau_r^n}$ son uniformemente integrables al ser S de clase D, vemos que

$$\lim_{r\to\infty} \sup_n \mathbb{E}\big(A_1^n \mathbf{1}_{A_1^n > r}\big) = 0.$$

$\square$

Gracias al Lema 4, podemos aplicar la Proposición 5.8 para garantizar la existencia de combinaciones convexas $\lambda_1^n, \ldots, \lambda_{N_n}^n$ con $N^n \geq n$ tales que

$$\lambda_n^n M_1^n + \cdots + \lambda_{N_n}^n M_1^{N_n}$$

converge en L_1 a una variable aleatoria M_1. Al definir $M_t = \mathbb{E}(\, M_1 \mid \mathscr{F}_t)$ (escogiendo una versión càdlàg) la desigualdad de Jensen nos permite ver que

$$\lambda_n^n M_t^n + \cdots + \lambda_{N_n}^n M_t^{N_n} \to M_t$$

en L_1 conforme $n \to \infty$. Para cada $n \geq 1$, extendemos a A^n a $[0,1]$ al hacerlo constante por pedazos en cada $(k/2^n, (k+1)/2^n]$ (para que sea continuo por la izquierda) y notamos que al definir

$$\tilde{A}^n = \lambda_n^n A^n + \cdots + \lambda_{N_n}^n A^{N_n},$$

así como al proceso cadlag

$$A = S - M$$

se tiene que $\tilde{A}_t^n \to A_t$ en L_1 conforme $n \to \infty$ para toda $t \in \mathscr{P} = \cup_n \mathscr{P}_n$. Así, $A_0 = 0$. Terminaremos con la demostración del Teorema 5.2 al probar que A es no-decreciente y predecible. Al tomar una subsucesión, podemos asumir que la convergencia de $\tilde{A}_t^n$ a A_t es casi segura para toda $t \in \mathscr{P}$. Puesto que $\tilde{A}^n$ es no-decreciente, vemos que A es no-decreciente en $\mathscr{P}$ y por continuidad a la derecha, concluimos que A es no-decreciente.

Finalmente, notemos que A^n y $\tilde{A}^n$ son continuos por la izquierda y adaptados, por lo que son predecible. Para mostrar que A es predecible, veremos que

$$\limsup_n \tilde{A}_t^n = A_t \tag{8}$$

para toda $t \in [0,1]$ casi seguramente. Ya lo hemos probado para toda $t \in \mathscr{P}$ y la monotonía de $\tilde{A}^n$ y A nos dice que

$$\limsup_n \tilde{A}_t^n \leq A_t \quad \text{y} \quad \limsup_n \tilde{A}_t^n = A_t \text{ si } A \text{ es continuo en } t.$$

Así, la ecuación (8) sólo podría ser falsa en las discontinuidades de A. Sin embargo, puesto que A tiene una cantidad finita de saltos de tamaño mayor a $1/k$ para cada

$k \in \mathbb{N}$, mismos que ocurren sucesivamente en tiempos de paro, vemos que para probar (8) basta ver que

$$\mathbb{E}\left(\limsup_n \tilde{A}^n_\tau\right) = \mathbb{E}(A_\tau)$$

para todo tiempo de paro τ.

Sea τ un tiempo de paro. Por el lema de Fatou, aplicado a $\tilde{A}^n_\tau \leq \tilde{A}^n_1$ (donde el lado derecho converge a A_1 en L_1), vemos que

$$\liminf_n \mathbb{E}(A^n_\tau) \leq \liminf_n \mathbb{E}\left(\tilde{A}^n_\tau\right) \leq \limsup_n \mathbb{E}\left(\tilde{A}^n_\tau\right) \leq \mathbb{E}\left(\limsup_n \tilde{A}^n \tau\right) \leq \mathbb{E}(A_\tau)\,.$$

Por lo tanto, basta probar que $\lim_n \mathbb{E}(A^n_\tau) = \mathbb{E}(A_\tau)$. Si $\sigma_n = \lceil \tau 2^n \rceil / 2^n$, entonces σ_n es un tiempo de paro que decrece a τ conforme $n \to \infty$. Al usar que S es de clase D y que sus trayectorias son càdlàg obtenemos $\mathbb{E}(S_{\sigma_n}) \to \mathbb{E}(S_\tau)$. Por otra parte:

$$\mathbb{E}(A^n_\tau) = \mathbb{E}\left(A^n_{\sigma_n}\right) = \mathbb{E}(S_{\sigma_n}) - \mathbb{E}(M_0) \to \mathbb{E}(S_\tau) - \mathbb{E}(M_0) = \mathbb{E}(A_\tau)\,.$$

5. La integral estocástica y sus propiedades

Bibliografía

[BSV12] Mathias Beiglböck, Walter Schachermayer, and Bezirgen Veliyev, *A short proof of the Doob-Meyer theorem*, Stochastic Process. Appl. **122** (2012), no. 4, 1204–1209. MR 2914749

[CY80] C. Cocozza and M. Yor, *Démonstration d'un théorème de F. Knight à l'aide de martingales exponentielles*, Seminar on Probability, XIV (Paris, 1978/1979) (French), Lecture Notes in Math., vol. 784, Springer, Berlin, 1980, pp. 496–499. MR 580150 (82h:60085)

[Dam65] K. È. Dambis, *On decomposition of continuous submartingales*, Teor. Verojatnost. i Primenen. **10** (1965), 438–448. MR 0202179 (34 #2052)

[Dav79] Burgess Davis, *Brownian motion and analytic functions*, Ann. Probab. **7** (1979), no. 6, 913–932. MR 548889 (80j:30001)

[Doo84] J. L. Doob, *Classical potential theory and its probabilistic counterpart*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 262, Springer-Verlag, New York, 1984. MR 731258 (85k:31001)

[DS65] Lester E. Dubins and Gideon Schwarz, *On continuous martingales*, Proc. Nat. Acad. Sci. U.S.A. **53** (1965), 913–916. MR 0178499 (31 #2756)

[EK46] P. Erdös and M. Kac, *On certain limit theorems of the theory of probability*, Bull. Amer. Math. Soc. **52** (1946), 292–302. MR 0015705 (7,459b)

[EK47] P. Erdös and M. Kac, *On the number of positive sums of independent random variables*, Bull. Amer. Math. Soc. **53** (1947), 1011–1020. MR 0023011 (9,292g)

[Itô87] Kiyosi Itô, *Differential equations determining a Markoff process*, Selected papers (Daniel W. Stroock and S.R.S Varadhan, eds.), Springer-Verlag, New York, 1987, Translated from the 1942 Japanese original.

[Kac49] M. Kac, *On distributions of certain Wiener functionals*, Trans. Amer. Math. Soc. **65** (1949), 1–13. MR 0027960 (10,383b)

[Kal02] Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR MR1876169 (2002m:60002)

[Kni71] Frank B. Knight, *A reduction of continuous square-integrable martingales to Brownian motion*, Martingales (Rep. Meeting, Oberwolfach, 1970), Springer, Berlin, 1971, pp. 19–31. Lecture Notes in Math., Vol. 190. MR 0370741 (51 #6967)

[KS91] Ioannis Karatzas and Steven E. Shreve, *Brownian motion and stochastic calculus*, second ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991. MR 1121940 (92h:60127)

[KW67] Hiroshi Kunita and Shinzo Watanabe, *On square integrable martingales*, Nagoya Math. J. **30** (1967), 209–245. MR 0217856 (36 #945)

[Lév39] Paul Lévy, *Sur certains processus stochastiques homogènes*, Compositio Math. **7** (1939), 283–339. MR 0000919 (1,150a)

[Mey62] P. A. Meyer, *A decomposition theorem for supermartingales*, Illinois J. Math. **6** (1962), 193–205. MR 0159359 (28 #2576)

[Mey63] P.-A. Meyer, *Decomposition of supermartingales: the uniqueness theorem*, Illinois J. Math. **7** (1963), 1–17. MR 0144382 (26 #1927)

[Par05] K. R. Parthasarathy, *Probability measures on metric spaces*, AMS Chelsea Publishing, Providence, RI, 2005, Reprint of the 1967 original. MR 2169627 (2006d:60004)

[PS78] Sidney C. Port and Charles J. Stone, *Brownian motion and classical potential theory*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1978, Probability and Mathematical Statistics. MR 0492329 (58 #11459)

[RW00] L. C. G. Rogers and David Williams, *Diffusions, Markov processes, and martingales. Vol. 2*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2000, Itô calculus, Reprint of the second (1994) edition. MR 1780932 (2001g:60189)

[RY99] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR MR1725357 (2000h:60050)

[Sim05] Barry Simon, *Functional integration and quantum physics*, second ed., AMS Chelsea Publishing, Providence, RI, 2005. MR 2105995 (2005f:81003)

[Ste01] J. Michael Steele, *Stochastic calculus and financial applications*, Applications of Mathematics (New York), vol. 45, Springer-Verlag, New York, 2001. MR 1783083 (2001i:60080)

[SW73] Tokuzo Shiga and Shinzo Watanabe, *Bessel diffusions as a one-parameter family of diffusion processes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **27** (1973), 37–46. MR 0368192 (51 #4433)

[WZ65] Eugene Wong and Moshe Zakai, *On the convergence of ordinary integrals to stochastic integrals*, Ann. Math. Statist. **36** (1965), 1560–1564. MR 0195142 (33 #3345)