Universidad Nacional Autónoma de México

Facultad de Ciencias

Difusiones entrelazadas

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
MATEMÁTICO

PRESENTA:
BRANDON ISRAEL GARCÍA FLORES

DIRECTOR DE TESIS:
DR. GERÓNIMO URIBE BRAVO

CIUDAD UNIVERSITARIA, CIUDAD DE MÉXICO, 2016

ÍNDICE GENERAL

AGRADECIMIENTOS

Agradezco profundamente a mis abuelos, a mi tío Francisco y a Meredith por apoyarme incondicionalmente y de distintas formas a lo largo de la licenciatura, así como a todos mis amigos por escucharme cuando tuviera algún tipo de dificultad.
También quisiera agradecer al Dr. Gerónimo por el tiempo dedicado para antenderme cuando tuviera dudas, así como por los consejos que me ha dado. Finalmente, quiero agradecer a mis sinodales por el tiempo dedicado para revisar el trabajo y por las sugerencias hechas, pues trascienden más allá de este trabajo.

INTRODUCCIÓN

El propósito de la presente tesis es desarrollar el tema de Difusiones entrelazadas, tratando de exponer de manera clara y detallada los resultados que aquí se presentan. El trabajo está principalmente basado en el artículo *Intertwinning diffusions and wave equations* de Suomik Pal y Mykhaelo Shkolnikov, aunque, para poder ofrecer argumentos más completos y cálculos con mayor detalle, se ha cambiado la estructura de la exposición y se han adaptado algunos de los argumentos (y de los enunciados de los teoremas) allí encontrados.
El concepto de entrelazamiento surge de manera natural al estudiar las propiedades de los procesos de Markov, en particular, al buscar condiciones bajo las cuales una función de un proceso de Markov es de nuevo de Markov. El Capítulo 1 da una breve discusión acerca de estas condiciones y de las ideas que motivaron la definición de entrelazamiento. Debido a que este capítulo está destinado a servir de introducción, la explicación de algunas de las nociones y de la notación usada ahí es diferida hasta el Capítulo 2.
El Capítulo 2 contiene algunas definiciones y teoremas citados frecuentemente a lo largo de la tesis. Si bien la ideología del trabajo es la de tratar de argumentar la mayoría de los enunciados hechos; debido a que se quiso conservar la fluidez y a que la mayoría de los resultados son de conocimiento general, en este capítulo se han omitido casi todas las demostraciones. Aún así, para beneficio del interesado, se ha proporcionado una referencia precisa acerca de dónde encontrar cada una de ellas.
Los Capítulos 3 y 4 desarrollan el tema principal, donde se define formalmente lo que es un entrelazamiento, tanto para semigrupos de transición como para difusiones homogéneas. En el Capítulo 3 se prueban los principales teoremas del texto. Finalmente, en el Capítulo 4 se proporcionan ejemplos de entrelazamientos y se demuestran algunas pequeñas generalizaciones o enuncian algunas propiedades.

CAPÍTULO 1

FUNCIONES DE PROCESOS DE MARKOV

Para motivar lo que se discutirá en el capítulo, se planteará lo siguiente: considérese un sistema dinámico discreto que se desarrolla de manera homogénea a través del tiempo y el cual evoluciona "sin memoria", para el cual podemos pensar que la posición actual depende exclusivamente de la posición anterior, por lo que podemos escribir $x_{n+1} = f(x_n)$, donde x_n es la posición al tiempo n. Si se considera que en dicho sistema existe algún tipo de perturbación aleatoria, entonces puede modelarse por una cadena de Markov $(X_n)_{n\in\mathbb{N}}$, donde ahora se tiene la relación $X_{n+1} = f(X_n, U_{n+1})$ para alguna sucesión de variables $(U_n)_{n\in\mathbb{N}}$ i.i.d. Se está interesado en medir alguna cantidad asociada a este sistema, la cual cambia también respecto al tiempo y cuyas mediciones forman otro proceso $(Y_n)_{n\in\mathbb{N}}$. Se asume que el proceso del sistema y el de mediciones están relacionados de manera que existe una función ϕ tal que para cada n, se tiene $\phi(X_n) = Y_n$ o, de manera más general, que existe una sucesión de variables aleatorias $(V_n)_{n\in\mathbb{N}}$ i.i.d. e independientes del proceso $(X_n)_{n\in\mathbb{N}}$ tales que $Y_n = \phi(X_n, V_n)$. Esta situación corresponde al caso en el que las mediciones no pueden ser efectuadas de manera precisa, por lo que la información recolectada incluye algún tipo de ruido. Para efectos de estimación, se quisiera saber si el proceso $(Y_n)_{n\in\mathbb{N}}$ es de nuevo una cadena Markov. Esto es equivalente a la existencia de una sucesión de variables aleatorias i.i.d. $(W_n)_{n\in\mathbb{N}}$ y de una función g tal que $Y_{n+1} = g(Y_n, W_{n+1})$. De existir tal función, se debe satisfacer la relación

$$\phi(f(X_n, U_{n+1}), V_{n+1}) = g(\phi(X_n, V_n), W_{n+1})$$

Esta igualdad expresa la idea de que una vez obtenida una muestra de X_n se puede obtener una muestra de Y_{n+1} de dos maneras: obteniendo primero una muestra de X_{n+1} según f y después una de Y_{n+1} según ϕ u obteniendo primero una muestra de Y_n con ϕ y luego una de Y_{n+1} a través de g. De manera gráfica, se tiene el siguiente diagrama

$$\begin{tikzcd} X_n \arrow[r, "f"] \arrow[d, "\phi"'] & X_{n+1} \arrow[d, "\phi"] \\ Y_n \arrow[r, "g"'] & Y_{n+1} \end{tikzcd}$$

Esto indica, de cierto modo, qué camino seguir en el problema de encontrar criterios suficientes para que el proceso $(Y_n)_{n\in\mathbb{N}}$ sea de Markov. En este capítulo solo se tratará de resolver esta cuestión cuando el proceso $(Y_n)_{n\in\mathbb{N}}$ depende solamente del proceso $(X_n)_{n\in\mathbb{N}}$, es decir, cuando $Y_n = \phi(X_n)$ para cada n.

Antes de dar una primer respuesta general, se continuará con el caso en el que X es una cadena de Markov con espacio de estados numerable (por lo que el espacio de estados de Y también se puede pensar numerable). En este caso, se quisiera demostrar que

$$P(Y_{n+1} = y_{n+1}|Y_0 = y_0, \ldots, Y_n = y_n) = P(Y_{n+1} = y_{n+1}|Y_n = y_n), \quad n \in \mathbb{N} \tag{1.1}$$

Usando la propiedad de Markov de la cadena X,

$$\begin{aligned} P(Y_{n+1} = y_{n+1}|Y_0 = y_0, \ldots, Y_n = y_n) &= \frac{P(Y_0 = y_0, \ldots, Y_{n+1} = y_{n+1})}{P(Y_0 = y_0, \ldots, Y_n = y_n)} \\ &= \frac{\sum_{x_1,\ldots,x_n} P(X_0 = x_0, \ldots, X_n = x_n, Y_{n+1} = y_{n+1})}{\sum_{x_1,\ldots,x_n} P(X_0 = x_0, \ldots, X_n = x_n)} \\ &= \frac{\sum_{x_1,\ldots,x_n} P(Y_{n+1} = y_{n+1}|X_n = x_n) p_{x_0,\ldots,x_n}}{\sum_{x_1,\ldots,x_n} p_{x_0,\ldots,x_n}} \end{aligned}$$

donde las sumas se extienden sobre todos los posibles valores de los x_i para los cuales $\phi(x_i) = y_i$ y $p_{x_0,\ldots,x_n} = P(X_0 = x_0, \ldots, X_n = x_n)$. De manera similar,

$$P(Y_{n+1} = y_{n+1}|Y_n = y_n) = \frac{\sum_{x_n} P(Y_{n+1} = y_{n+1}|X_n = x_n)P(X = x_n)}{\sum_{x_n} P(X_n = x_n)}$$

Con el fin de obtener una igualdad entre estos cocientes, se puede imponer la condición

$$P(Y_{n+1} = y_{n+1}|X_n = x_n) = P(Y_{n+1} = y_{n+1}|X_n = x'_n) \text{ siempre y cuando } \phi(x_n) = \phi(x'_n) = y_n. \tag{1.2}$$

Asumiendo esto, si $x \in \phi^{-1}(y_n)$ es fijo, entonces

$$\begin{aligned}
P(Y_{n+1} = y_{n+1} | Y_0 = y_0, \ldots, Y_n = y_n) &= \frac{\sum_{x_1,\ldots,x_n} P(Y_{n+1} = y_{n+1} | X_n = x_n) p_{x_0,\ldots,x_n}}{\sum_{x_1,\ldots,x_n} p_{x_0,\ldots,x_n}} \\
&= \frac{\sum_{x_1,\ldots,x_n} P(Y_{n+1} = y_{n+1} | X_n = x) p_{x_0,\ldots,x_n}}{\sum_{x_1,\ldots,x_n} p_{x_0,\ldots,x_n}} \\
&= \frac{P(Y_{n+1} = y_{n+1} | X_n = x) \sum_{x_1,\ldots,x_n} p_{x_0,\ldots,x_n}}{\sum_{x_1,\ldots,x_n} p_{x_0,\ldots,x_n}} \\
&= P(Y_{n+1} = y_{n+1} | X_n = x)
\end{aligned}$$

y del mismo modo, $P(Y_{n+1} = y_{n+1} | Y_n = y_n) = P(Y_{n+1} = y_{n+1} | X_n = x)$, por lo que (1.1) se cumple y Y es cadena de Markov.
Como se puede observar, lo único que hizo falta para obtener una respuesta en este caso particular fue la ecuación (1.2). Generalizando esta condición se obtiene una primer respuesta a la pregunta planteada:

Teorema 1.1. *[Criterio de Dynkin] Sea X un proceso de Markov con semigrupo de transición $(P_t)_{t\geq 0}$ y espacio de estados $(E, \mathscr{E})$. Asúmase que $\phi : (E, \mathscr{E}) \to (F, \mathscr{F})$ es medible y satisface*

$$\phi(A) \in \mathscr{F} \quad \text{para todo } A \in \mathscr{E}. \tag{1.3}$$

Si, más aún, para cada $C \in \mathscr{F}$ se cumple

$$P_t(x, \phi^{-1}(C)) = P_t(x', \phi^{-1}(C)) \quad \text{si} \quad \phi(x) = \phi(x') \tag{1.4}$$

entonces el proceso $Y = \phi(X)$ es de nuevo un proceso de Markov bajo P_ν para cualquier medida inicial ν.

Demostración. Para $t \geq 0$, defínase $Q_t : F \times \mathscr{F} \to [0, 1]$ por

$$Q_t(y, C) = \begin{cases} P_t(x, \phi^{-1}(C)) & \text{si } y = \phi(x) \text{ para algún } x \in E, \\ 0 & \text{si } y \notin \phi(E) \end{cases}$$

Se afirma que $(Q_t)_{t\geq 0}$ es un semigrupo de transición. Por (1.4), Q_t está bien definida y es claro que $Q_t(y, \cdot)$ es una medida, pues es la medida constante 0 o la medida imagen de $P_t(x, \cdot)$ bajo ϕ para alguna $x \in E$. Dado $D \in \mathscr{F}$, se tiene

$$\begin{aligned}
Q_t(\cdot, C)^{-1}(D) &= [Q_t(\cdot, C)^{-1}(D) \cap \phi(E)] \cup [Q_t(\cdot, C)^{-1}(D) \setminus \phi(E)] \\
&= \phi(P_t(\cdot, \phi^{-1}(C))^{-1}(D)) \cup [Q_t(\cdot, C)^{-1}(D) \setminus \phi(E)]
\end{aligned}$$

Como ϕ satisface la ecuación (1.3) y $P_t(\cdot,\phi^{-1}(C))$ es medible, se tiene $\phi(P_t(\cdot,\phi^{-1}(C))^{-1}(D)) \in \mathscr{F}$ y debido a que $Q_t(\cdot,C)^{-1}(D) \setminus \phi(E)$ siempre es igual a $\emptyset$ o a $F \setminus \phi(E)$, se concluye que $Q_t(\cdot,C)$ es medible para todo $C \in \mathscr{F}$. Así, si $y = \phi(x)$,

$$\begin{aligned} Q_{t+s}(y,C) &= P_{t+s}(x,\phi^{-1}(C)) \\ &= \int_E P_t(x,dy)P_s(y,\phi^{-1}(A)) \\ &= \int_E P_t(x,dy)Q_s(\phi(y),A) \\ &= \int_F Q_t(y,dy')Q_s(y',A) \end{aligned}$$

donde en la última igualdad se ha hecho un cambio de variable, usando de nueva cuenta que $Q_t(y,\cdot)$ es la medida imagen de $P_t(x,\cdot)$ bajo ϕ. Se concluye que $Q_{t+s} = Q_tQ_s$. Para cualquier $x \in E$, la igualdad $Q_t(\phi(x),C) = P_t(x,\phi^{-1}(C))$, puede escribirse como $(Q_t1_C)\circ\phi = P_t(1_A\circ\phi)$. Por linealidad y el método de aproximación estándar, se sigue que para cualquier $f : E \to \mathbb{R}^+$ medible,

$$(Q_tf)\circ\phi = P_t(f\circ\phi). \tag{1.5}$$

De este modo,

$$\begin{aligned} P_\nu(f(\phi(X_t))|\sigma(X_u,u \le s)) &= P_{t-s}(f\circ\phi)(X_s) \\ &= [(Q_{t-s}f)\circ\phi](X_s) \\ &= (Q_{t-s}f)(\phi(X_s)) \end{aligned}$$

por lo que Y es un proceso de Markov respecto a la filtración $(\sigma(X_u,u \le t))_{t\ge0}$ y en particular, respecto a la filtración inducida por él mismo. □

La prueba hace ver que solo es necesario comprobar que (1.3) se satisface solo para los conjuntos de la forma $P_t(\cdot,\phi^{-1}(C))^{-1}(D)$ con $C, D \in \mathscr{F}$. Usando esto, se probará que dado un movimiento browniano d-dimensional, $(X_t)_{t\ge0}$, si $\phi : \mathbb{R}^d \to [0,\infty)$ es la función norma, entonces el proceso definido por $Y_t = \phi(X_t) = |X_t|$, para $t \ge 0$ es un proceso de Markov (el cual es el conocido proceso de Bessel d-dimensional). Aquí, $F = [0,\infty)$ y $\mathscr{F} = \mathscr{B}(F)$ y consideramos a F inmerso en $\mathbb{R}^d$ a través de la inclusión en la primer coordenada. Dado C un boreliano en F, $\phi^{-1}(C)$ es una unión de esferas $d-1$ dimensionales centradas en el origen que pasan por cada punto en C, por lo que $\phi^{-1}(C)$ es invariante bajo rotaciones. El semigrupo de transición del movimiento browniano $(P_t)_{t\ge0}$ tiene densidades definidas por $p_t(x,z) = (2\pi t)^{-1/2}\exp(-|x-z|^2/(2t)))$, por lo que si $|x| = |y|$ para algún $y \in \mathbb{R}^d$, se tiene que probar que

$$\int_{\phi^{-1}(C)} (2\pi t)^{-1/2}\exp(-|x-z|^2/(2t)))dz = \int_{\phi^{-1}(C)} (2\pi t)^{-1/2}\exp(-|y-z|^2/(2t)))dz \tag{1.6}$$

Sea $R : \mathbb{R}^d \to \mathbb{R}^d$ una rotación tal que $R(y) = x$ (la cual existe porque $|x| = |y|$). Como R^{-1} también es una rotación, entonces $R^{-1}(\phi^{-1}(C)) = \phi^{-1}(C)$. Al hacer un cambio de variable, se tiene

$$\begin{aligned}\int_{\phi^{-1}(C)} (2\pi t)^{-1/2} \exp(-|x-z|^2/(2t)))dz &= \int_{R^{-1}(\phi^{-1}(C))} (2\pi t)^{-1/2} \exp(-|x-R(z)|^2/(2t)))dz \\ &= \int_{\phi^{-1}(C)} (2\pi t)^{-1/2} \exp(-|R(y-z)|^2/(2t)))dz \\ &= \int_{\phi^{-1}(C)} (2\pi t)^{-1/2} \exp(-|y-z|^2/(2t)))dz\end{aligned}$$

Lo cual muestra que en este caso se cumple (1.4). Ahora, (1.4) implica que si $x \in P_t(\cdot, \phi^{-1}(C))^{-1}(B)$ para algún boreliano $B \in \mathscr{B}(E)$, entonces $\phi^{-1}(\phi(x)) \subset P_t(\cdot, \phi^{-1}(C))^{-1}(B)$. Como $\phi^{-1}(\phi(x))$ es una $d-1$-esfera, el conjunto $P_t(\cdot, \phi^{-1}(C))^{-1}(B)$ es también una unión de esferas $d-1$ dimensionales centradas en el origen. Luego, la imagen bajo ϕ de $P_t(\cdot, \phi^{-1}(C))^{-1}(B)$ es simplemente su intersección con $[0, \infty)$, por lo que $\phi(P_t(\cdot, \phi^{-1}(C))^{-1}(B)) \in \mathscr{B}(E)$. Se sigue entonces que el proceso de Bessel es de Markov.

Aunque el Teorema 1.1 da una primer respuesta a la pregunta planteada, ésta está alejada de ser satisfactoria, pues aún para espacios de estados y funciones ϕ con propiedades agradables, las hipótesis del teorema pueden no ser satisfechas. Por ejemplo, considérese $E = \mathbb{R}^2$, $F = \mathbb{R}$, ambos con sus σ-álgebras de Borel y ϕ la proyección en la primera entrada. Aunque en este ejemplo ϕ es infinitamente diferenciable y suprayectiva, se sabe que existe un conjunto $A \in \mathscr{B}(\mathbb{R}^2)$ tal que $\phi(A) \notin \mathscr{B}(\mathbb{R})$, por lo que incluso en esta situación, el teorema no puede ser utilizado.

Otra dificultad es que en general, podría ser complicado comprobar las hipótesis del teorema para funciones arbitrarias ϕ, pues para verificar (1.3) hace falta hacerlo para todos los conjuntos $A \in \mathscr{E}$ (o al menos los de la forma $P_t(\cdot, \phi^{-1}(C))^{-1}(B)$), pues, contrario a la imagen inversa, la imagen directa no se comporta bien respecto a la intersección de conjuntos (en el sentido de que, en general, no se tiene $\phi(A \cap B) = \phi(A) \cap \phi(B)$). Así, no es posible restringirse a clases más pequeñas de conjuntos y usar teoremas de extensión como el de las clases monótonas o el de los $\lambda - \pi$ sistemas.

Para tratar de remediar esto, obsérvese que la propiedad que permite finalizar la prueba es la ecuación (1.5). Si Φ es el núcleo de transición de E a F inducido por ϕ (de modo que para una función medible acotada $f : F \to \mathbb{R}$, $\Phi f = f \circ \phi$), entonces, debido a (1.5), el teorema se puede plantear de la siguiente manera: si existe una colección de núcleos de transición $(Q_t)_{t\geq 0}$ en F de manera que

$$P_t \Phi = \Phi Q_t, \qquad t \geq 0 \tag{1.7}$$

entonces Y es un proceso de Markov. La ecuación (1.7) muestra que cierta condición de “entrelazamiento” entre los núcleos es suficiente para obtener una respuesta satisfactoria.

Continuando en esta dirección, se pueden obtener aún más condiciones de este tipo que garanticen que Y es de nuevo Markov: la probabilidad condicional $P(Y_t \in C \mid \sigma(Y_u, u \leq t))$ define una función que depende solo de Y_t y de C. Sin embargo, la condición $Y_t \in C$ es equivalente a la condición $X_t \in A$ con $A = \phi^{-1}(C)$, por lo que se puede pensar que dicha función podría depender solo de Y_t y de A. Es decir, se debería tener $P(X_t \in A \mid \sigma(Y_u, u \leq t)) = L(Y_t, A)$, donde, por las propiedades de la probabilidad condicional, L debería también definir un núcleo de F a E. Así, si Y ha de ser un proceso de Markov de nuevo, sus probabilidades de transición deben estar dadas por

$$\begin{aligned} Q_t f(Y_0) &= \mathbb{E}[f(Y_t) \mid Y_0] \\ &= \mathbb{E}[f \circ \phi(X_t) \mid X_0 \mid Y_0]^1 \\ &= \mathbb{E}[P_t(f \circ \phi)(X_0) \mid Y_0] \\ &= LP_t(f \circ \phi)(Y_0). \end{aligned}$$

Usando notación de composición de núcleos, esto es equivalente a $Q_t = LP_t\Phi$ con Φ definida como antes. Aunque esto solo indica qué condiciones deberían ser satisfechas, hacen falta más condiciones para garantizar que las probabilidades de transición $(Q_t)_{t\geq 0}$ forman un semigrupo. Antes de enunciar un teorema que resuelve esto, se enunciará un lema que será requerido para la demostración.

Lema 1.2. *Un proceso estocástico X con valores en $(E, \mathscr{E})$ es un proceso de Markov con respecto a la filtración $(\mathscr{F}_t)_{t\geq 0}$ donde $\mathscr{F}_t = \sigma(X_u, u \leq t)$, probabilidades de transición $(P_{s,t})_{t,s\geq 0}$ y medida inicial ν si y solo si para cualesquiera $0 = t_0 < t_1 < \cdots < t_n$ y $f_0, f_1, \ldots, f_n$ funciones medibles no negativas,*

$$\mathbb{E}\Big[\prod_{k=0}^{n} f_k(X_{t_k})\Big] = \int_E \nu(dx_0) f_0(x_0) \int_E P_{0,t_1}(x_0, dx_1) f_1(x_1) \cdots \int_E P_{t_{n-1}.t_n}(x_{n-1}, dx_n) f_n(x_n).$$

Teorema 1.3. *Supóngase que existe un núcleo de transición L de F a E tal que*

1. *$L\Phi = I$, el núcleo de transición identidad en F, es decir, para cualquier función medible acotada f, $L(\Phi(f)) = f$[2,3].*

2. *para cada $t \geq 0$, el núcleo de transición $Q_t = LP_t\Phi$ de F a F satisface la identidad $LP_t = Q_t L$.*

Si X tiene distribución inicial $\lambda = L(y, \cdot)$, donde $y \in F$, entonces Y es un proceso de Markov con semigrupo de transición $(Q_t)_{t\geq 0}$ y estado inicial y (en el sentido de que $P_\lambda(Y_0 = y) = 1$).

[1] Aquí y en adelante, para una variable aleatoria Z no negativa o acotada y dos sigma álgebras $\mathscr{F}$ y $\mathscr{G}$, la esperanza condicional iterada $\mathbb{E}[\mathbb{E}[Z|\mathscr{G}]\,|\mathscr{F}]$ se denotará por $\mathbb{E}[Z|\mathscr{G}|\mathscr{F}]$. Una notación similar se usará al condicionar respecto a variables aleatorias.

[3] Véase el Capítulo 2, Definición **??**

Demostración. Por la condición 2, se tiene

$$Q_{t+s} = LP_{t+s}\Phi = LP_tP_s\Phi = Q_tLP_s\Phi = Q_tQ_s$$

y del mismo modo se ve que $Q_{t+s} = Q_sQ_t$, por lo que la familia $(Q_t)_{t\geq 0}$ es un semigrupo.

Si $C \in \mathscr{F}$, la condición 1 implica que

$$1_C = L\Phi 1_C = L(1_C \circ \phi) = L1_{\phi^{-1}(C)} = L(\cdot, \phi^{-1}(C)) \tag{1.8}$$

por lo que $\phi^{-1}(C)$ tiene medida cero respecto a $L(y, \cdot)$ si $y \notin C$ o la medida del espacio en caso contrario. Así, si $g : E \to \mathbb{R}$ es acotada y medible,

$$\begin{aligned} L[(\Phi 1_C)g](y) = L[(1_C \circ \phi)g](y) = \int_E 1_{\phi^{-1}(C)}(y)g(y)L(y, dy) &= \int_{\phi^{-1}(C)} g(y)L(y, dy) \\ &= \begin{cases} \int_E g(y)L(x, dy) & y \in C \\ 0 & y \notin C \end{cases} \\ &= 1_C(y)\int_E g(y)L(y, dy) = (1_CLg)(y) \end{aligned}$$

de donde se obtiene $L[(\Phi 1_C)g] = 1_CLg$. Por linealidad, se sigue que $L[(\Phi f)g] = fLg$ para cualquier función simple medible $f : F \to \mathbb{R}$ y usando el teorema de convergencia dominada, se sigue que

$$L[(\Phi f)g] = fLg$$

para cualquier función medible acotada f, lo que aunado a la condición 2 implica

$$LP_t[(\Phi f)g] = Q_tL[(\Phi f)g] = Q_t[fLg], \qquad t \geq 0. \tag{1.9}$$

Se afirma que para cualesquiera $t_1, \ldots, t_n \geq 0$, $f_1, \ldots, f_n$ funciones medibles acotadas en F y g medible acotada en E, se tiene

$$LP_{t_1}[(\Phi f_1)P_{t_2}[(\Phi f_2)\cdots P_{t_n}[(\Phi f_n)g]]] = Q_{t_1}[f_1Q_{t_2}[f_2\cdots Q_{t_n}[f_nLg]]]. \tag{1.10}$$

Procediendo por inducción, el caso base se demostró en (1.9), por lo que se asume que el resultado es cierto para algún natural n. Así, para $t_1, \ldots, t_{n+1} \geq 0$ y $f_1, \ldots, f_{n+1}$ funciones medibles acotadas en F y g medible acotada en E, se tiene que $h = P_{t_{n+1}}[(\Phi f_{n+1})g]$ es una función medible acotada en E, por lo que la hipótesis de inducción y (1.9) implican que

$$\begin{aligned} LP_{t_1}[(\Phi f_1)P_{t_2}[(\Phi f_2)\cdots P_{t_{n+1}}[(\Phi f_{n+1})g]]] &= LP_{t_1}[(\Phi f_1)P_{t_2}[(\Phi f_2)\cdots P_{t_n}[(\Phi f_n)h]]] \\ &= Q_{t_1}[f_1Q_{t_2}[f_2\cdots Q_{t_n}[f_nLh]]] \\ &= Q_{t_1}[f_1Q_{t_2}[f_2\cdots Q_{t_n}[f_nL[P_{t_{n+1}}[(\Phi f_{n+1})g]]]]] \\ &= Q_{t_1}[f_1Q_{t_2}[f_2\cdots Q_{t_{n+1}}[f_{n+1}Lg]]]. \end{aligned}$$

Así, el lema y (1.10) con $g = 1$ implican que para $0 = t_0 < t_1 < \cdots < t_n$ y $f_0, f_1 \ldots, f_n$ funciones medibles acotadas en F,

$$\begin{aligned}
\mathbb{E}_\lambda\Big[\prod_{k=0}^{n} f_k(Y_{t_k})\Big] &= \mathbb{E}_\lambda\Big[\prod_{k=0}^{n} f_k \circ \phi(X_{t_k})\Big] \\
&= \int_E L(y, dx_0) f_0 \circ \phi(x_0) \int_E P_{t_1}(x_0, dx_1) f_1 \circ \phi(x_1) \cdots \int_E P_{t_n - t_{n-1}}(x_{n-1}, dx_n) f_n \circ \phi(x_n) \\
&= L P_0[(\Phi f_0) P_{t_1}[(\Phi f_1) \cdots P_{t_n}[(\Phi f_n) g]]](y) \\
&= Q_0[f_0 Q_{t_1}[f_1 \cdots Q_{t_n}[f_n L g]]](y) \\
&= f(y) \int_F Q_{t_1}(y, dy_1) f_1(y_1) \cdots \int_F Q_{t_n}(y_{n-1}, dy_n) f_n(y_n)
\end{aligned}$$

Lo cual, por el lema 1.2, implica que Y es un proceso de Markov (homogéneo) con semigrupo de transición $(Q_t)_{t \geq 0}$ y estado inicial y. □

Observaciones.

1. La condición 1 del teorema implica dos cosas: que ϕ es una función suprayectiva y, asumiendo que los conjuntos unitarios están en $\mathscr{F}$, que $L(y, \phi^{-1}(\{y\}) = 1$. Ambas afirmaciones son consecuencias de la ecuación (1.8), pues, para la primera, ésta implica que si $\phi^{-1}(C) = \phi^{-1}(D)$, entonces $1_C = 1_D$, por lo que $C = D$. En particular, como $\phi^{-1}(\phi(E)) = \phi^{-1}(F) = E$, se tiene $\phi(E) = F$. La segunda afirmación es clara tomando $C = \{y\}$, pues así $L(y, \phi^{-1}(\{y\})) = 1_{\{y\}}(y) = 1$. Se ve entonces que la condición 1 es natural si se espera que L actúe como distribución condicional de X_t dado Y_t.

2. Debido a la discusión que precede al teorema, cuando la condición 1 se cumple, la condición 2 es equivalente a $P(X_t \in A \mid \sigma(Y_u, u \leq t)) = L(Y_t, A)$, lo cual se sigue de la ecuación (1.9).

Podemos observar que el punto 2 del teorema 1.3 muestra de nuevo una relación similar a la de la ecuación (1.7). Este tipo de relaciones son los que dan origen al concepto de *entrelazamiento* y los resultados expuestos hasta ahora muestran que la relación tiene interés en sí misma. Los entrelazamientos y sus principales consecuencias serán desarrollados en los Capítulos 3 y 4 en donde se ha de restringir al contexto de las difusiones para poder utilizar más herramientas de la teoría de la probabilidad.

CAPÍTULO 2

DEFINICIONES Y RESULTADOS PRELIMINARES

Se comenzará dando algunas definiciones y resultados sin demostración, los cuales se usarán a lo largo del trabajo. Las pruebas de éstos pueden hallarse en [2], [4], [7].

Procesos de Feller y generador infinitesimal

Definición 2.1. Dados dos espacios medibles $(E, \mathscr{E})$ y $(F, \mathscr{F})$, un núcleo o probabilidad de transición de E a F es una función $N : E \times \mathscr{F} \to [0, 1]$ tal que para cada $x \in E$, la función $A \mapsto N(x, A)$ es una medida de probabilidad en F y para cada $A \in \mathscr{F}$ fijo, la función $x \mapsto N(x, A)$ es una función medible en E.

Dado un núcleo de transición N de E a F se puede definir un único operador lineal continuo $\tilde{N}$ de las funcions medibles acotadas en F a las funciones medibles acotadas en E dado por

$$(\tilde{N}f)(x) = \int_F f(y)N(x, dy).$$

Este operador es no negativo, en el sentido de que $\tilde{N}f \geq 0$ si $f \geq 0$; satisface $N1_F = 1_E$ y al ver los espacios como espacios de Banach con la norma del supremo esencial, $\tilde{N}$ es también continuo. Si al contrario se inicia con un operador lineal continuo no negativo M de las funciones medibles y acotadas en E a las funciones medibles y acotadas en F, se puede definir un núcleo de transición $\widehat{M}$ por la fórmula

$$\widehat{M}(x, A) = (M1_A)(x)$$

y este núcleo es tal que $M = \tilde{\widehat{M}}$. Debido a esta relación, en todo el trabajo cualquier núcleo de transición se considera también un operador lineal entre espacios de funciones medibles acotadas y no se hace distinción entre la notación de ambas nociones. Un *semigrupo de transición* es entonces una colección de núcleos de transición $(P_t)_{t\geq 0}$ tales que para cada $t, s \geq 0$, $P_{t+s} = P_tP_s$, donde, para una función adecuada f, la composición de núcleos de transición se calcula como $(P_tP_s)(f) = P_t(P_s(f))$.
A partir de este momento y a lo largo de la sección se dará por hecho que se tiene un espacio métrico separable y localmente compacto E, el cual se piensa también como un espacio medible al equiparlo con su σ-álgebra de Borel $\mathscr{E} = \mathscr{B}(E)$. Denótese por $C_0 = C_0(E)$ a la clase de funciones

continuas $f : E \to \mathbb{R}$ tales que $f(x) \to 0$ cuando $x \to \infty$, donde esto significa que para todo $\varepsilon > 0$ existe un conjunto compacto $K \subset E$ tal que $|f(x)| < \varepsilon$ cuando $x \in E \setminus K$. El espacio C_0 puede convertirse en un espacio de Banach al dotarlo de la norma $\|f\| = \sup_{x\in E} |f(x)|$. Usando esto, se puede dar la siguiente definición, donde un operador positivo es un operador que manda funciones postivas en funciones postivas.

Definición 2.2. Un semigrupo de Feller en $C_0(E)$ es una familia $(T_t)_{t\geq 0}$ de operadores lineales positivos en $C_0(E)$ tal que

i) $T_0 = \mathrm{Id}$ y $\|T_t\| \leq 1$ para cada t,

ii) $T_{t+s} = T_t T_s$ para cualesquiera $s, t \geq 0$ y

iii) $\lim_{t\downarrow 0} \|T_t f - f\| = 0$ para cada $f \in C_0(E)$

Cuando las condiciones (i) y (ii) son satisfechas por un semigrupo $(T_t)_{t\geq 0}$, la última condición es de hecho equivalente a la siguiente

iii') $\lim_{t\downarrow 0} T_t f(x) = f(x)$ para cada $f \in C_0$ y $x \in E$.

Con cada semigrupo de Feller en $C_0(E)$ se puede asociar una única función de transición homogénea $(P_t)_{t\geq 0}$ en $(E, \mathscr{E})$ tal que $T_t f(x) = P_t f(x)$ para cada $f \in C_0$ y $x \in E$. Una función de transición asociada a un semigrupo de Feller es una *función de transición de Feller*. Un proceso de Markov cuya función de transición es de Feller es llamado un *proceso de Feller*.
Aunque los procesos de Feller poseen varias características, aquí solo se estará interesado en la definición y propiedades de su generador, las cuales se describen enseguida.

Definición 2.3. Sea X un proceso de Feller. Se dice que una función $f \in C_0$ pertenece al dominio $\mathcal{D}$ del generador infinitesimal de X si el límite

$$\mathcal{A}f = \lim_{t\downarrow 0} \frac{1}{t}(P_t f - f) \tag{2.1}$$

existe en C_0. El operador $\mathcal{A} : \mathcal{D} \to C_0$ así definido es llamado el *generador infinitesimal* del proceso X o del semigrupo P_t.

Si el rol del dominio $\mathcal{D}$ se quiere enfatizar, se dirá que el proceso X (o su semigrupo) tiene generador $(\mathcal{A}, \mathcal{D})$. Se procede entonces a dar una lista de propiedades del generador $\mathcal{A}$.

Teorema 2.4. *Sea $(P_t)_{t\geq 0}$ un semigrupo de Feller con generador $(\mathcal{A}, \mathcal{D})$. Entonces, $T_t f \in \mathcal{D}$ para toda $f \in \mathcal{D}$ y se satisface*

$$T_t f - f = \int_0^t T_s \mathcal{A} f ds, \qquad f \in \mathcal{D},\ t \geq 0$$

Más aún, $T_t f$ es diferenciable en 0 si y solo si $f \in \mathcal{D}$, en cuyo caso

$$\frac{d}{dt}(T_t f) = T_t \mathcal{A} f = \mathcal{A} T_t f, \quad t \geq 0$$

Demostración. Consúltese [4], página 372, Teorema 17.6. □

En la definición usual del generador infinitesimal, se pide que el semigrupo $(T_t)_{t\geq 0}$ sea de Feller, sin embargo, esta condición puede ser eliminida, en cuyo caso se tiene que sustituir C_0 por el conjunto

$$B_0 = \{f \in L_\infty(E) \mid \lim_{t\to 0} \|T_t f - f\| = 0\}$$

Y, en este caso, el dominio del generador se extiende a las funciones $f \in B_0$ para las cuales el límite en (2.1) existe y resultados análogos a los de esta sección se siguen cumpliendo (véase [7], página 291). Se entenderá, pues, que este es el dominio del generador cuando se hable de él (aún cuando el semigrupo se de Feller) y si se requiere considerar solo las funciones en C_0, esto se hará explícito.

Integración estocástica

Dada una semimartingala continua X, la variación cuadrática de X será denotada por $\langle X, X\rangle$. Si K es un proceso progresivamente medible y localmente acotado, la integral de K respecto a X se denotará indistintamente por

$$(K \cdot X) = \int_0^\cdot K_s dX_s$$

Esto abarca el caso en el que X es solamente una martingala local o un proceso de variación finita, por lo que si $X = M + A$ para M una martingala local y A un proceso de variación finita, entonces

$$K \cdot X = K \cdot M + K \cdot A$$

Si $X = (X^1, \ldots, X^d)$ es una martingala d-dimensional y $K = (K^1, \ldots, K^d)$ es un proceso para el cual cada entrada es progresivamente medible y localmente acotada, entonces se escribe

$$\int_0^\cdot K_s dX_s = \sum_{i=1}^d \int_0^\cdot K_s^i X_s^i$$

Con esta notación, se tiene el siguiente

Teorema 2.5 (Dambis, Dubins-Schwarz)**.** *Si M es una martingala local continua respecto a la filtración derecha por la continua $(\mathscr{F}_t)_{t\geq 0}$ tal que $M_0 = 0$ y tal que $\langle M, M\rangle_\infty = \infty$ y si se define*

$$T_t = \inf\{s : \langle M, M\rangle_s > t\}$$

entonces $B_t = M_{T_t}$ es un movimiento browniano respecto a la filtración $(\mathscr{F}_{T_t})_{t\geq 0}$ y $M_t = B_{\langle M,M\rangle_t}$.

Demostración. Consúltese [7], página 181, Teorema 1.6. □

Procesos de Itô y difusiones

En esta sección, a y b denotarán funciones a las matrices de tamaño $d \times d$ y a $\mathbb{R}^d$, respectivamente, sujetas a las condiciones

i) a y b son Borel medibles y localmente acotadas,

ii) para cada x, $a(x)$ es simétrica y no negativa definida, esto es, para cualquier $\lambda \in \mathbb{R}^d$, $\sum_{i,j} a_{ij}(x)\lambda_i\lambda_j \geq 0$.

Con cada par (a, b) se asocia un operador diferencial de segundo orden

$$\mathcal{A} = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(\cdot)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(\cdot)\frac{\partial}{\partial x_i}$$

Así, si C_0^∞ denota la clase de funciones infinitamente diferenciables en $\mathbb{R}^d$ con soporte acotado, se tiene

Definición 2.6. Un proceso de Markov X en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathscr{F}_t, P, P_x)$ con espacio de estados $\mathbb{R}^d$ es una *difusión* con generador $\mathcal{A}$ si

i) X tiene trayectorias continuas y

ii) Para cada $x \in \mathbb{R}^d$ y cualquier $f \in C_0^\infty$,

$$\mathbb{E}_x[f(X_t)] = f(x) + \mathbb{E}_x\left[\int_0^t \mathcal{A}f(X_s)ds\right]$$

En esta situación, se dice que X tiene *covarianza* o *coeficiente de difusión* a y *deriva* b.

Obsérvese que también se podría dejar que las funciones a y b dependan del tiempo, por lo cual se obtendría, para cada tiempo s, un operador diferencial elíptico de segundo orden

$$\mathcal{A}_s = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(s,\cdot)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(s,\cdot)\frac{\partial}{\partial x_i}$$

Esto extendería la definición de difusión a la de *difusión no homogénea.* En este caso, se tendrían medidas de probabilidad $P_{s,x}$ correspondientes al proceso iniciado en x al tiempo s y se requeriría que para cada $f \in C_0^\infty$, $s < t$,

$$\mathbb{E}_{s,x}[f(X_t)] = f(x) + \mathbb{E}_{s,x}\left[\int_s^t \mathcal{A}_u f(X_u)du\right]$$

Extendiendo esta definición, se puede omitir la condición de ser proceso de Markov, obteniendo la siguiente

Definición 2.7. Sea β un movimiento Browniano r-dimensional respecto a la filtración $(\mathscr{F}_t)$ (denotado $(\mathscr{F}_t) - BM^r$) definido en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ y σ (respectivamente, b) un proceso predecible localmente acotado con valores en las matrices de tamaño $d \times r$ (respectivamente, $\mathbb{R}^d$). Decimos que un proceso continuo y adaptado X es un proceso de Itô con covarianza $\sigma\sigma^t$ y deriva b si X satisface la ecuación

$$X_t = X_0 + \int_0^t \sigma(s) d\beta_s + \int_0^t b(s) ds.$$

Si en particular se tiene dos funciones $\tilde{\sigma}$ y $\tilde{b}$ definidas en $\mathbb{R}^d$ tales que $\sigma(s) = \tilde{\sigma}(X_s)$ y $b(s) = \tilde{b}(X_s)$ y $(X_t, P_x, x \in \mathbb{R}^d)$ es un proceso de Markov, la fórmula de Itô implica entonces que, para cualquier función $f \in C_0^\infty$,

$$f(X_t) = f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dX_s^i + \frac{1}{2} \int_0^t \sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) \sum_{k=1}^d \tilde{\sigma}_{ik}(X_s) \tilde{\sigma}_{jk}(X_s) ds$$

Al definir $\tilde{a} = \tilde{\sigma}\tilde{\sigma}^t$, se tiene que $\tilde{a}_{ij} = \sum_{k=1}^d \tilde{\sigma}_{ik} \tilde{\sigma}_{kj}^t = \sum_{k=1}^d \tilde{\sigma}_{ik} \tilde{\sigma}_{jk}$, por lo que

$$\begin{aligned} \mathbb{E}_x[f(X_t)] &= f(x) + \mathbb{E}_x \left[\int_0^t \sum_{i=1}^d \frac{\partial f}{\partial x_i}(X_s) \tilde{b}(X_s) + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) \sum_{k=1}^d \tilde{\sigma}_{ik}(X_s) \tilde{\sigma}_{jk}(X_s) ds \right] \\ &= f(x) + \mathbb{E}_x \left[\int_0^t \mathcal{A} f(X_s) ds \right] \end{aligned}$$

donde $\mathcal{A} = \frac{1}{2} \sum_{i,j=1}^d \tilde{a}_{ij}(\cdot) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d \tilde{b}_i(\cdot) \frac{\partial}{\partial x_i}$, lo que justifica que los procesos de Itô sean extensiones de las difusiones.

Pensado en el proceso canónico, un proceso de Markov con trayectorias continuas, $(X_t)_{t \geq 0}$, es una difusión si y solo si para cualquier función $f \in C_0^\infty$, el proceso $(M_t^f)_{t \geq 0}$ definido por

$$M_t^f = f(X_t) - f(X_0) - \int_0^t \mathcal{A} f(X_s) ds$$

es una martingala para cualquier $\mathbb{E}_x$. En efecto, para ver esto solo hace falta ver que esto es equivalente a la condición $ii)$ de la definición 2.6. Si M_t^f es una martingala y $x \in \mathbb{R}^d$, entonces, como $M_0^f = 0$,

$$0 = \mathbb{E}_x[M_0^f] = \mathbb{E}_x[M_t^f] = \mathbb{E}_x \left[f(X_t) - f(X_0) - \int_0^t \mathcal{A} f(X_s) ds \right] = \mathbb{E}_x[f(X_t)] - f(x) - \left[\int_0^t \mathcal{A} f(X_s) ds \right]$$

que es la condición $ii)$. Recíprocamente, si dicha condición es la que se cumple, entonces, para cualesquiera $0 \leq s, t$, se tiene $\mathbb{E}_{X_s}[M_t^f] = 0$ (pues es $\mathbb{E}_x[M_t^f] = 0$ para toda $x \in \mathbb{R}^d$). Luego, por la propiedad de Markov, si $s \leq t$,

$$\begin{aligned}\mathbb{E}_x[M_t^f \mid \mathscr{F}_s] &= \mathbb{E}_x[M_s^f + M_t^f - M_s^f \mid \mathscr{F}_s]\\ &= M_s^f + \mathbb{E}_x\left[f(X_t) - f(X_s) - \int_s^t \mathcal{A}f(X_u)du \;\middle|\; \mathscr{F}_s \right]\\ &= M_s^f + \mathbb{E}_{X_s}[M_{t-s}^f] = M_s^f\end{aligned}$$

por lo que $(M_t^f)_{t\geq 0}$ es una martingala. Esta observación lleva a la siguiente definición.

Definición 2.8. Una medida de probabilidad P en W es una solución al *Problema de la Martingala relativo a* $\mathcal{A} = \frac{1}{2}\sum_{i,j=1}^d a_{ij}(\cdot)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(\cdot)\frac{\partial}{\partial x_i}$ si

i) $P[X_0 = x] = 1$

ii) para cualquier $f \in C_0^\infty$, el proceso

$$M_t^f = f(X_t) - f(X_0) - \int_0^t \mathcal{A}f(X_s)ds$$

es una P-martingala con respecto a la filtración $(\sigma(X_s, s \leq t))$.
Se dice que el Problema de la martingala está *bien planteado* si para cada x existe exactamente una solución.

En el contexto de la definición (2.7), si nuevamente se tiene $\sigma(s) = \tilde{\sigma}(X_s)$ y $b(s) = \tilde{b}(X_s)$ para dos funciones $\tilde{\sigma}$ y $\tilde{b}$ definidas en $\mathbb{R}^d$ y si $X_0 = x$ c.s., entonces la medida imagen de P bajo X, $X(P)$, es una solución al problema de la martingala $\pi(x, \sigma\sigma^t, b)$.
Los últimos resultados de esta sección dan condiciones para saber cuándo el problema de la martingala está bien planteado y si es que al obtener una solución para cada x, el conjunto de soluciones está relacionado de manera que el proceso canónico es una difusión con $\mathcal{A}$ como su generador.

Teorema 2.9. *Si el problema de la martingala $\pi(x, a, b)$ está bien planteado y si para cada $A \in \mathscr{B}(\mathbb{R}^d)$ y $t \geq 0$, la aplicación $x \to P_x[X_t \in A]$ es medible, entonces $(X_t, P_x, x \in \mathbb{R}^d)$ es un proceso de Markov con función de transición $P_t(x, A) = P_x[X_t \in A]$, donde (X_t) es el proceso canónico.*

Demostración. Consúltese [7], página 371, Teorema 1.9. □

Teorema 2.10. *Sea $(\Omega, \mathscr{F}_t, P)$ un espacio de probabilidad filtrado, $W = C(\mathbb{R}_+, \mathbb{R}^d)$ y f y g dos funciones predecibles acotadas a las matrices de tamaño $d \times r$ y a $\mathbb{R}^d$ respectivamente. Asúmase que existe una constante K tal que para cada t y w,*

$$|f(t, w) - f(t, w')| + |g(t, w) - g(t, w')| \leq K \sup_{s \leq t} |w(s) - w'(s)|.$$

Entonces, existe un único proceso X_t^x, $x \in \mathbb{R}^d$, $t \in \mathbb{R}_+$ con trayectorias continuas con respecto a ambas variables x y t, tal que para cada x,

$$X_t^x = x + \int_0^t f(s, X_\cdot^t) dB_s + \int_0^t g(s, X_\cdot^x) ds \quad P - \text{c.s.}$$

donde B_t es un movimiento $(\mathscr{F}_t)$-Browniano estándar.

Demostración. Consúltese [7], página 379, Teorema 2.4. □

El sentido de unicidad en este teorema es en el sentido de que si existe otro proceso Y_t^x que satisfaga lo mismo, entonces Y_t^x y X_t^x son indistinguibles.
Obsérvese que bajo las condiciones del Teorema (2.10), el problema de la martingala está bien definido, donde la solución está dada por $P_x = X^x(P)$. Si Y_t es el proceso canónico, entonces, por continuidad, las aplicaciones $x \to P_x[Y_t \in A]$ son medibles para cada $A \in \mathscr{B}(\mathbb{R}^d)$ y $t \geq 0$. Así, usando los Teoremas (2.9) y (2.10), se tiene

Teorema 2.11. *Sean σ y b funciones Lipschitz continuas de $\mathbb{R}^d$ a las matrices de tamaño $d \times d$ y $\mathbb{R}^d$ respectivamente y defínanse f y g en $\mathbb{R}_+ \times W$ por $f(s,w) = \sigma(w(s))$ y $g(s,w) = b(w(s))$. Sea X_t^x el proceso del Teorema (2.10). Defínase a la función de transición P_t por $P_t(x, A) = P[X_t^x \in A]$. Entonces, la función de transición P_t es de Feller.*

Demostración. Consúltese [7], página 380, Teorema 2.5. □

Resultados de análisis funcional

Aquí se pensará que B es un espacio de Banach sobre el campo $\mathbb{K}$, el cual puede ser $\mathbb{R}$ o $\mathbb{C}$.

Definición 2.12. Un operador lineal $\mathcal{A}$ con dominio $\mathcal{D}$ en B es llamado *cerrado* si satisface alguna de las siguientes propiedades equivalentes

i) Para cualquier sucesión $(f_n) \subset \mathcal{D}$ tal que $f_n \to f$ y $\mathcal{A}f_n \to g$ para ciertos elementos $f, g \in B$ se tiene $f \in \mathcal{D}$ y $g = \mathcal{A}f$.

ii) La *gráfica* $G = \{(f, Af); f \in \mathcal{D}\}$ es un subconjunto cerrado de $B \times B$.

iii) El espacio $B_1 = (\mathcal{D}, \|\cdot\|_{\mathcal{A}})$ es un espacio de Banach para la norma de la gráfica

$$\|x\|_{\mathcal{A}} := \|x\| + \|\mathcal{A}x\|, x \in \mathcal{D}$$

En general, se dice que un operador $\mathcal{A}$ es un operador *cerrable* si la cerradura de su gráfica $\overline{G}$ es la gráfica de un operador $\overline{\mathcal{A}}$, el cual es llamado la *cerradura* de $\mathcal{A}$. Obsérvese que $\mathcal{A}$ es cerrable si y solo si las condiciones $f_n \in \mathcal{D}$, $f_n \to 0$ y $\mathcal{A}f_n \to g$ implican $g = 0$.
Cuando $\mathcal{A}$ es cerrado, un *centro* para $\mathcal{A}$ está definido como un subespacio lineal $D \subset \mathcal{D}$ tal que la restricción $\mathcal{A}|_D$ tiene cerradura $\mathcal{A}$. En relación la primer sección, se tiene lo siguiente

Teorema 2.13. *El generador $\mathcal{A}$ de un semigrupo de Feller es cerrado y para cualquier $\lambda > 0$, un subespacio $D \subset \mathcal{D}(\mathcal{A})$ es un centro para $\mathcal{A}$ si y solo si $(\lambda - \mathcal{A})D$ es denso en C_0.*

Demostración. Consúltese [4], página 374, Lema 19.8. $\square$

En el caso especial cuando $S = \mathbb{R}^d$, se dice que un operador $\mathcal{A}$ tal que $C_0^\infty \subset \mathrm{dom}(\mathcal{A})$ es *local* en C_0^∞ si cada vez que f se anule en una vecindad de $x \in \mathbb{R}^d$ se tiene $\mathcal{A}f(x) = 0$.
Regresando al caso general, se tiene lo siguiente

Definición 2.14. Sea $B^* = \{y^* : B \to \mathbb{K} \mid y^*\text{es continua}\}$. Para un operador densamente definido $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ en B, se define el *operador adjunto* $(\mathcal{A}^*, \mathcal{D}(\mathcal{A}^*))$ en B^* por

$$
\begin{aligned}
D(\mathcal{A}^*) &= \{x^* \in B^* \mid \text{ existe } y^* \in B^* \text{ tal que } x^*\mathcal{A}x = y^*x \text{ para todo } x \in \mathcal{D}(\mathcal{A})\} \\
\mathcal{A}^*x^* &= y^* \text{ para } x^* \in \mathcal{D}(\mathcal{A}^*)
\end{aligned}
$$

Para enunciar el teorema siguiente, se requerirá de la definición de derivada en un espacio de Banach y del problema de Cauchy abstracto. Dada una función $u : \mathbb{R} \to B$, se dice que u es *diferenciable* en t si el límite

$$
\frac{d}{dt}u(t) = \dot{u}(t) = \lim_{h \to 0} \frac{u(t+h) - u(t)}{h}
$$

existe en B. A la función $\dot{u}$ así definida se le llama la *derivada* de u.
El problema del valor inicial

$$
\begin{cases} \dot{u}(t) = \mathcal{A}u(t) & t \geq 0 \\ u(0) = x \end{cases} \tag{2.2}
$$

es llamado el *problema de Cauchy abstracto* asociado a $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ y el valor inicial x. Una función $u : \mathbb{R}_+ \to B$ es llamada una *solución clásica* del problema de Cauchy abstracto si u es continuamente diferenciable, $u(t) \in \mathcal{D}(\mathcal{A})$ para toda $t \geq 0$ y la ecuación ((2.2)) se cumple.

Teorema 2.15. *Sea $(\mathcal{A}, \mathcal{D})$ el generador de un semigrupo de Feller $(T_t)_{t\geq 0}$. Entonces, para cada $x \in \mathcal{D}$, la función*

$$
u : t \mapsto u(t) := T_t x
$$

es la única solución clásica al problema de Cauchy abstracto.

Demostración. Consúltese [2], página 145, Proposición 6.2. $\square$

CAPÍTULO 3

DIFUSIONES ENTRELAZADAS

En este capítulo se expondrán de manera detallada las pruebas de los Teoremas 1 y 2 de [6], para lo cual se comenzará dando algunas definiciones previas.
Como se vio en el primer capítulo, al intentar responder la pregunta de cuándo una función de un proceso de Markov es de nuevo un proceso de Markov, se obtienen relaciones de la forma $Q_t L = L P_t$, donde $(Q_t)_{t\geq 0}$ y $(P_t)_{t\geq 0}$ son semigrupos de transición y L es un núcleo de transición. La primer definición formaliza esta relación, obteniendo un lo que es conocido como un *entrelazamiento*.

Definición 3.1. Sean $(Q_t)_{t\geq 0}$, $(P_t)_{t\geq 0}$ dos semigrupos de transición en los espacios medibles $(E_1, \mathscr{E}_1)$, $(E_2, \mathscr{E}_2)$, respectivamente. Supóngase que L es un núcleo de transición que manda funciones acotadas en $\mathscr{E}_2$ a aquellas en $\mathscr{E}_1$. Se dice que el par ordenado (Q, P) está entrelazado con enlace L si para cada $t \geq 0$ la relación $Q_t L = L P_t$ se cumple (donde se considera que ambos lados son operadores actuando en funciones medibles acotadas en $\mathscr{E}_2$). En este caso, se denota $Q\langle L\rangle P$.

En lo subsecuente, se pensará que X y Y son dos difusiones homogéneas con semigrupos de transición $(P_t)_{t\geq 0}$ y $(Q_t)_{t\geq 0}$ y espacios de estados $\mathcal{X} \subset \mathbb{R}^m$ y $\mathcal{Y} \subset \mathbb{R}^n$.

Definición 3.2. Una difusión $Z = (Z^1, Z^2)$ con espacio de estados $\mathcal{X} \times \mathcal{Y}$ es llamada un entrelazamiento de las difusiones X y Y con enlace L si las condiciones siguientes se cumplen

(i) Z^1 y Z^2 tienen la misma distribución que X y Y respectivamente y

$$\mathbb{E}[f(Z_0^1) \mid Z_0^2 = y] = (Lf)(y), \tag{3.1}$$

para todas las funciones f medibles acotadas en $\mathcal{X}$.

(ii) Los semigrupos de transición P y Q están entrelazados, $Q\langle L\rangle P$ (por el inciso anterior, el semigrupo de Z^1 es también P y el semigrupo de Z^2 es Q).

(iii) El proceso Z^1 es de Markov respecto a la filtración generada por (Z^1, Z^2).

(iv) Para cualquier $t \geq 0$, dada Z_t^2, la variable aleatoria Z_t^1 es condicionalmente independiente de $\sigma(Z_s^2, s \leq t)$ y se distribuye condicionalmente de acuerdo a L. Esto es, para cualquier conjunto medible A,

$$P[Z_t^1 \in A \mid \sigma(Z_s^2, s \leq t)] = P[Z_t^1 \in A \mid Z_t^2] = L(Z_t^2, A).$$

(v) Para cualquier $t \geq 0$, dadas Z_t^1 y Z_0^2, las variables aleatorias Z_0^1 y Z_t^2 son condicionalmente independientes.

En este caso, se escribe $Z = Y\langle L\rangle X$.

En la ecuación (3.1) se entiende que Lf es una de las funciones para las cuales se cumple que $(Lf)(Z_0^2) = \mathbb{E}[f(Z_0^1) \mid Z_0^2]$. Supóngase que para un proceso $(U_t)_{t\geq 0}$, $U_t = (U_t^1, U_t^2)$ existe algún proceso $(\eta_t)_{t\geq 0}$ independiente de $(U_t^1)_{t\geq 0}$ se tiene que U_t^2 es $\sigma(\mathscr{F}_t, \mathscr{G}_t)$ medible, donde $(\mathscr{F}_t)_{t\geq 0}$ es la filtración generada por U^1 y $(\mathscr{G}_t)_{t\geq 0}$ es la filtración generada por η. Si U^1 es de Markov con semigrupo $(T_t)_{t\geq 0}$, se sigue entonces, que para cualquier función f en $\mathcal{X}$ medible y acotada y $0 \leq s, t$,

$$
\begin{aligned}
\mathbb{E}[f(U_{t+s}^1) \mid \sigma(U_u, 0 \leq u \leq s)] &= \mathbb{E}[f(U_{t+s}^1) \mid \sigma(\mathscr{F}_s, \mathscr{G}_s) \mid \sigma(U_u, 0 \leq u \leq s)] \\
&= \mathbb{E}[f(U_{t+s}^1) \mid \mathscr{F}_s \mid \sigma(U_u, 0 \leq u \leq s)] \\
&= \mathbb{E}[(T_t f)(U_s^1) \mid \sigma(U_u, 0 \leq u \leq s)] = (T_t f)(U_s^1)
\end{aligned}
$$

por lo que U^1 es de Markov respecto a la filtración generada por (U^1, U^2). Así, la condición (iii) captura la idea de que Z^2 se obtiene a partir de Z^1 con posible ruido independiente. La condición (iv) se entiende mejor en el marco de un diagrama parecido al presentado en el primer capítulo

$$
\begin{tikzcd}
Z_n^2 \arrow[r, "Q_s"] \arrow[d, "L"'] & Z_{t+s}^2 \arrow[d, "L"] \\
Z_t^1 \arrow[r, "P_s"'] & Z_{t+s}^1
\end{tikzcd}
$$

Las flechas verticales están invertidas respecto a aquel diagrama, pues aquí se representa la acción que tienen las variables aleatorias respecto a las medidas. De este modo, dada una muestra de Z_t^2 se puede condicionar según Q_s para obtener una muestra de Z_{t+s}^2 y luego condicionar a este valor para obtener una muestra de Z_{t+s}^1 según L; o, usando la propiedad de entrelazamiento, este camino debería ser equivalente a obtener un valor de Z_t^1 según L y condicionar a este nuevo valor para obtener una muestra de Z_{t+s}^1 según P_s. La condición (iii) implica que el proceso $(Z_t^2, Z_t^1, Z_{t+s}^1)_{t\geq 0}$ es de Markov. La condición (iv) se pide entonces para que el proceso $(Z_t^2, Z_{t+s}^2, Z_{t+s}^1)_{t\geq 0}$ sea también de Markov. Finalmente, la condición (v) solo se usa para asegurar que la distribución de conjunta de Z_0 y Z_t está únicamente determinada por las condiciones (iii) y (iv).

Considérese dos generadores infinitesimales

$$\mathcal{A}^X = \sum_{i=1}^{m} b_i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{m} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j}$$

$$\mathcal{A}^Y = \sum_{k=1}^{n} \gamma_k \frac{\partial}{\partial y_k} + \frac{1}{2} \sum_{k,l=1}^{n} \rho_{kl} \frac{\partial^2}{\partial y_k \partial y_l}$$

donde $b = (b_i)$ es una función en $\mathbb{R}^m$ continua en el interior de $\mathcal{X}$, $\gamma = (\gamma_k)$ es una función en $\mathbb{R}^n$ continua en el interior de $\mathcal{Y}$ y $a = (a_{ij})$ y $\rho = (\rho_{kl})$ son matrices positivas definidas y simétricas de tamaños $m \times m$ y $n \times n$ continuas en los interiores de $\mathcal{X}$ y $\mathcal{Y}$ respectivamente. Se supone además que estos generadores cumplen las siguientes dos condiciones

Condición 1. El problema de la martingala relativo a $\mathcal{A}^X$ en $\mathcal{X}$ está bien planteado en el sentido de la Definición 2.9 y su solución es un proceso de Feller. Además que el conjunto de funciones $C_0^\infty(\mathcal{X})$ es un centro para $\mathcal{A}^X$. Se supone lo análogo para $\mathcal{A}^Y$.

Condición 2. Se consideran las siguiente condiciones de regularidad para el núcleo L:

- L está definido como un operador integral

  $$(Lf)(y) = \int_{\mathcal{X}} \Lambda(y,x) f(x) dx$$

  mandando $C_0(\mathcal{X})$ en $C_0(\mathcal{Y})$.

- $\Lambda(\cdot, x)$ es estrictamente positiva y continuamente diferenciable en $\mathcal{Y}$ para cada $x \in \mathcal{X}$ fijo. Se define $V = \log \Lambda$ y se denota por $\nabla_y V$ al gradiente de V con respecto a y, por lo que

  $$(\nabla_y V)(\cdot, x) = \frac{1}{\Lambda} \Big(\frac{\partial}{\partial y_1} \Lambda(\cdot, x), \ldots, \frac{\partial}{\partial y_n} \Lambda(\cdot, x) \Big)$$

Para un punto $z \in \mathbb{R}^{m+n}$ escríbase $z = (x, y)$ con $x \in \mathbb{R}^m$ y $\mathbb{R}^n$. Si f es una función de dos variables, $(x, y) \mapsto f(x, y)$, f_x denotará la función obtenida al fijar la primer variable y similarmente se define f_y.

Teorema 3.3. *Sean X y Y dos difusiones dadas como las soluciones de los problemas de la martingala relativos a $\mathcal{A}^X$ y $\mathcal{A}^Y$ respectivamente. Sea $Z = (Z^1, Z^2)$ una difusión en $\mathcal{X} \times \mathcal{Y}$ cuyo generador está dado por*

$$\mathcal{A}^Z = \mathcal{A}^X + \mathcal{A}^Y + (\nabla_y V)' \rho \nabla_y \tag{3.2}$$

y supóngase que $\mathcal{A}^Z$ satisface la condicion 1. Más aún, supóngase que la condición inicial de Z satisface

$$P(Z_0^1 \in B \mid Z_0^2 = y) = \int_B \Lambda(y, x) dx, \quad B \in \mathscr{B}(\mathbb{R}^m)$$

Si Λ es tal que

(i) Λ_x *está en el dominio de* $\mathcal{A}^Y$ *en* $C_0(\mathcal{Y})$ *para todo* $x \in \mathcal{X}$ *y* $\mathcal{A}^Y\Lambda$ *es continua en* $\mathcal{Y} \times \mathcal{X}$ *y acotada en* $\mathcal{Y} \times K$ *para cualquier* $K \subset \mathcal{X}$ *compacto,*

(ii) la medida L_y *pertenece al dominio de* $(\mathcal{A}^X)^*$ *para todo* $y \in \mathcal{Y}$ *(donde se ve al dual de* $C_0(\mathcal{X})$ *como el espacio de medidas regulares en* $\mathcal{X}$*) y*

(iii) la medida $[(\mathcal{A}^X)^* L_y]$ *es absolutamente continua con respecto a la medida de Lebesgue y su densidad está dada por* $(\mathcal{A}^Y\Lambda_x)$*.*

entonces, $Z = Y\langle L\rangle X$

La última condición sobre Λ implica que, para toda $f \in \mathcal{D}(\mathcal{A}^X)$ y $y \in \mathcal{Y}$,

$$\begin{aligned} \int_{\mathcal{X}} (\mathcal{A}^X f)(x)\Lambda(y,x)dx &= \int_{\mathcal{X}} (\mathcal{A}^X f)(x)L(y,dx) \\ &= \int_{\mathcal{X}} f(x)[(\mathcal{A}^X)^* L_y](dx) \\ &= \int_{\mathcal{X}} (\mathcal{A}^Y\Lambda_x)(y)f(x)dx \end{aligned} \tag{3.3}$$

y de la definición del adjunto, este cálculo muestra que la dicha condición es equivalente a esta igualdad. Al tomar la restricción a un subespacio más pequeño (por ejemplo, $C_0^\infty(\mathcal{X})$), $\mathcal{A}^X$ es un operador simétrico, por lo que la medida $(\mathcal{A}^X)^* L_y$ tiene densidad $\mathcal{A}^X\Lambda_y$ y la condición (iii) del teorema se puede escribir como

$$\mathcal{A}^X\Lambda_y = \mathcal{A}^Y\Lambda_x.$$

En general, esta es una ecuación diferencial parcial hiperbólica y en el caso en el que X y Y son movimientos brownianos, se obtiene la ecuación de onda unidimensional clásica. Esto conduce a un primer ejemplo de la aplicación de este teorema: se hallará un entrelazamiento entre $X = B_1 + Z_0^1$ y $Y = B_2 + Z_0^2$, donde B_1 y B_2 son dos moviemientos brownianos independientes y (Z_0^1, Z_0^2) son variables aleatorias que satisfacen la ecuación (3.1). Obsérvese que lo generadores de X y Y son $\frac{\partial^2}{\partial x^2}$ y $\frac{\partial^2}{\partial y^2}$ respectivamente, los cuales satisfacen la Condición 1. Considérese el núcleo

$$\Lambda(y,x) = \frac{1}{\pi(1+(y-x)^2)}$$

Obsérvese que $\mathcal{A}^Y\Lambda(y,x) = \frac{6(y-x)^2-2}{\pi(1+(y-x)^2)^3}$. Así, $\mathcal{A}^Y\Lambda_x \in C_0$ y $\Lambda_x \in \mathcal{D}(\mathcal{A}^Y)$. Como $\frac{\partial^2}{\partial x^2}\Lambda(x,y) = \frac{\partial^2}{\partial y^2}\Lambda(y,x)$, al integrar por partes se obtiene la ecuación (3.3).
Defínase el operador

$$\begin{aligned} \mathcal{A}^Z &= (\nabla_y V)\rho\nabla_y + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \\ &= -\frac{2(y-x)}{1+(y-x)^2}\frac{\partial}{\partial y} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \end{aligned}$$

Sean $\sigma : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ y $b : \mathbb{R}^2 \to \mathbb{R}^2$ dadas por $\sigma(x, y) = \mathbb{I}$ y $b(x, y) = (0, -\frac{2(y-x)}{1+(y-x)^2})$, con $\mathbb{I}$ la matriz identidad. Se tiene entonces que σ y b son funciones Lipschitz continuas, por lo que el Teorema (3.3) implica que $\mathcal{A}^Z$ satisface la Condición 1 y existe un proceso de Feller $Z = (Z^1, Z^2)$ con condición inicial (Z^1_0, Z^2_0) y generador $\mathcal{A}^Z$, que, por el Teorema 3.3, es un entrelazamiento de X y Y.

Algunos lemas se requerirán para la prueba del Teorema 3.3. De aquí en adelante, se pensará que el espacio de probabilidad sobre el que se trabaja es el espacio canónico de trayectorias continuas $C([0, \infty), \mathcal{X} \times \mathcal{Y})$ con la σ-álgebra (producto) de Borel, $\mathscr{F}$, y con medida de probabilidad P dada por la distribución de Z. En el espacio se define la filtración continua por la derecha $(\mathscr{F}_t)_{t\geq 0}$ generada por las proyecciones y completada respecto a la familia de medidas $(P_z)_{z\in\mathcal{X}\times\mathcal{Y}}$, el conjunto de soluciones al problema de la martingala para $\mathcal{A}^Z$ iniciando en $z \in \mathcal{X} \times \mathcal{Y}$. También se definen las filtraciones $(\mathscr{F}^X_t)_{t\geq 0}$ y $(\mathscr{F}^Y_t)_{t\geq 0}$ definidas como las filtraciones continuas por la derecha y completas (respecto a (P_z)) generadas por las primeras m y las últimas n proyecciones respectivamente. Obsérvese que estas filtraciones son subfiltraciones de $(\mathscr{F}_t)_{t\geq 0}$.

Lema 3.4. *Para cada función $f \in C^\infty_0(\mathcal{X} \times \mathcal{Y})$, la función $\Lambda_x f_x$ pertenece al dominio de $\mathcal{A}^Y$ y se satisface*

$$\mathcal{A}^Y(\Lambda_x f_x) = \Lambda_x(\mathcal{A}^Y f_x) + (\nabla_y\Lambda)'\rho(\nabla_y f) + f_x(\mathcal{A}^Y\Lambda_x) \tag{3.4}$$

Si $f \in C^\infty_0(\mathcal{Y})$, se cumple la misma ecuación con f en lugar de f_x.

Demostración. Solo se probará el caso en el que $f \in C^\infty_0(\mathcal{X} \times \mathcal{Y})$, pues la prueba es análoga. Sea $x \in \mathcal{X}$ y $\tilde{\Lambda}_x$ una función continua en $\mathbb{R}^n$ con soporte compacto contenido en el interior de $\mathcal{Y}$, la cual, además, coincide con Λ_x en una vecindad abierta del soporte de f_x. Sea ϕ una función de densidad en $\mathbb{R}^n$ con soporte compacto e infinitamente diferenciable. Para cada $q \in \mathbb{N}$, sea ϕ_q dada por $\phi_q(y) = q^n\phi(qy)$. Las funciones ϕ_q tienen las siguientes dos propiedades

i) El diámetro del soporte de ϕ_q tiende a cero cuando q tiende a infinito y

ii) Si $g : \mathbb{R}^n \to \mathbb{R}$ es una función continua, la sucesión de funciones $(g_q)_{q\in\mathbb{N}}$ dada por

$$g_q(y) = (f * \phi_q)(y) = \int_{\mathbb{R}^n} g(y - \tilde{y})\phi_q(\tilde{y})d\tilde{y} = \int_{\mathbb{R}^n} g(\tilde{y})\phi_q(y - \tilde{y})d\tilde{y}$$

es infinitamente diferenciable y converge a g puntualmente. Si además, g es uniformemente continua en un conjunto, la convergencia es uniforme sobre este. En particular, la convergencia es uniforme en conjuntos compactos.

Para cada $q \in \mathbb{N}$, sea $\varphi_{q,x}(y) = \tilde{\Lambda}_x * \phi_q$. Por la propiedad (i) de la sucesión $(\phi_q)_{q\in\mathbb{N}}$, el soporte de $\varphi_{q,x}$ eventualmente estará contenido en el interior de $\mathcal{Y}$, por lo que $\varphi_{q,x} \in C_0(\mathcal{Y})$ para q suficientemente grande. Además, $\frac{\partial}{\partial y_i}\varphi_{q,x}(y) = (\frac{\partial}{\partial y_i}\tilde{\Lambda}_x) * \phi_q$, $i \in \{1, \ldots, n\}$, de donde

$$\varphi_{q,x} \to \tilde{\Lambda}_x \text{ y } \nabla\varphi_{q,x} \to \nabla\tilde{\Lambda}_{q,x} \text{ uniformemente en el soporte de } f_x. \tag{3.5}$$

Usando que C_0^∞ es un centro para $\mathcal{A}^Y$, existe una sucesión de funciones $(\Lambda_{r,x})_{r\in\mathbb{N}}$ que converge a Λ_x uniformemente y tales que la sucesión $(\mathcal{A}^Y\Lambda_{q,x})_{r\in\mathbb{N}}$ converge a $\mathcal{A}^Y\Lambda_x$ uniformemente. Así, diferenciando bajo la integral, para cada q se tiene $\mathcal{A}^Y(\Lambda_{r,x}*g) = (\mathcal{A}^Y\Lambda_{r,x})*g$ donde g es cualquier función en $C_0^\infty(\mathcal{Y})$. El lado derecho de esta igualdad converge uniformemente a $(\mathcal{A}^Y\Lambda_x)*g$, por lo que, al usar que A es cerrado, se obtiene

$$\mathcal{A}^Y(\Lambda_x * g) = (\mathcal{A}^Y\Lambda_x)*g, \quad g \in C_0^\infty(\mathcal{Y})$$

Luego, como A es un operador local y $\tilde{\Lambda}_x = \Lambda_x$ en una vecindad abierta del soporte de f_x, para cada $q \in \mathbb{N}$ y y en dicha vecindad, se tiene $\mathcal{A}^Y(\tilde{\Lambda}_x * \phi_q)(y) = \mathcal{A}^Y(\Lambda_x * \phi_q)(y)$. Así,

$$\mathcal{A}^Y\varphi_{q,x} = \mathcal{A}^Y(\tilde{\Lambda}_x * \phi_q) = \mathcal{A}^Y(\Lambda_x * \phi_q) = (\mathcal{A}^Y\Lambda_x)*\phi_q \to \mathcal{A}^Y\Lambda_x \text{ uniformemente en el soporte de } f_x. \tag{3.6}$$

Ahora, como cada $\varphi_{q,x}f_x$ es infinitamente diferenciable con soporte compacto, cada una de ellas está en $\mathcal{D}(\mathcal{A}^Y)$, por lo que la regla de Leibniz implica que

$$\mathcal{A}^Y(\varphi_{q,x}f_x) = \varphi_{q,x}(\mathcal{A}^Y f_x) + (\nabla_y\varphi_{q,x})'\rho(\nabla_y f) + f_x(\mathcal{A}^Y\varphi_{q,x}) \tag{3.7}$$

Fuera del soporte de f_x, cada uno de los sumandos en el lado derecho de (3.7) se anula, por lo que (3.5) y (3.6) implican que

$$\mathcal{A}^Y(\varphi_{q,x}f_x) \to \Lambda_x(\mathcal{A}^Y f_x) + (\nabla_y\Lambda)'\rho(\nabla_y f) + f_x(\mathcal{A}^Y\Lambda_x) \text{ uniformemente en } \mathcal{Y} \tag{3.8}$$

Por (3.5), $(\varphi_{q,x}f_x)_{q\in\mathbb{N}}$ converge uniformemente a $\Lambda_x f_x$ en $\mathcal{Y}$, por lo que la ecuación (3.8) y el que $\mathcal{A}^Y$ sea cerrado implican que $\Lambda_x f_x \in \mathcal{D}(\mathcal{A}^Y)$ y

$$\mathcal{A}^Y(\Lambda_x f_x) = \Lambda_x(\mathcal{A}^Y f_x) + (\nabla_y\Lambda)'\rho(\nabla_y f) + f_x(\mathcal{A}^Y\Lambda_x)$$

como se quería. □

Lema 3.5. *Si $f \in C_0^\infty(\mathcal{X}\times\mathcal{Y})$, entonces la función $F : \mathcal{Y} \to \mathbb{R}$ dada por $y \mapsto (Lf_y)(y)$ pertenece al dominio de $\mathcal{A}^Y$ y*

$$\mathcal{A}^Y F(y) = \int_{\mathcal{X}} \mathcal{A}^Y(\Lambda_x f_x)(y)dx \tag{3.9}$$

Si $f \in C_0^\infty(\mathcal{X})$, la función Lf pertenece al dominio de $\mathcal{A}^Y$ y

$$\mathcal{A}^Y(Lf)(y) = \int_{\mathcal{X}} (\mathcal{A}^Y\Lambda_x)(y)f(x)dx \tag{3.10}$$

Demostración. Primero se probará el caso en el que $f \in C_0^\infty(\mathcal{X}\times\mathcal{Y})$. Sea K el soporte de f y $K_{\mathcal{X}}$ la proyección de K en $\mathcal{X}$, por lo que $K_{\mathcal{X}}$ es compacto. Para cada $n \in \mathbb{N}$ sea $\{\mathcal{X}_r^n\}_{1\le r\le R_n}$

una partición finita de K_X en conjuntos medibles con diámetro menor a n^{-1} y sea $x_r^n \in \mathcal{X}_r^n$, $r = 1, \ldots, R_n$. Defínase la sucesión de funciones $F_n : \mathcal{Y} \to \mathbb{R}$ dada por

$$F_n = \sum_{r=1}^{R_n} \Lambda_{x_r^n} f_{x_r^n} m(\mathcal{X}_r^n)$$

donde m es la medida de Lebesgue en $\mathbb{R}^m$. Obsérvese que para cada n, $F_n \in \mathcal{A}^Y$ y

$$\mathcal{A}^Y F_n = \sum_{r=1}^{R_n} \mathcal{A}^Y (\Lambda_{x_r^n} f_{x_r^n}) m(\mathcal{X}_r^n).$$

Ahora, como f se anula fuera de K, las funciones Λf y $\mathcal{A}^Y(\Lambda.f.)$ son uniformemente continuas en $\mathcal{X} \times \mathcal{Y}$) (recuérdese que $\mathcal{A}^Y$ es un operador local), por lo que para cada $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que para todo $n \geq N$ y $(x_1, y), (x_2, y) \in \mathcal{X} \times \mathcal{Y}$ tal que $|x_1 - x_2| < n^{-1}$, se tiene

$$|\Lambda(y, x_1) f(x_1, y) - \Lambda(y, x_2) f(x_2, y)| < \varepsilon \quad \text{y} \quad |\mathcal{A}^Y(\Lambda_{x_1} f_{x_1})(y) - \mathcal{A}^Y(\Lambda_{x_2} f_{x_2})(y)| < \varepsilon \tag{3.11}$$

Por lo tanto,

$$\begin{aligned} |F(y) - F_n(y)| &= \left| \sum_{r=1}^{R_n} \int_{\mathcal{X}_r^n} (\Lambda(y, x) f(x, y) - \Lambda(y, x_r^n) f(x_r^n, y)) dx \right| \\ &\leq \sum_{r=1}^{R_n} \int_{\mathcal{X}_r^n} |\Lambda(y, x) f(x, y) - \Lambda(y, x_r^n) f(x_r^n, y)| dx \\ &\leq \sum_{r=1}^{R_n} \int_{\mathcal{X}_r^n} \varepsilon dx = \varepsilon m(K_{\mathcal{X}}) \end{aligned}$$

De donde se obtiene que F_n converge uniformemente a F y, de manera similar, se ve que $\mathcal{A}^Y F_n$ converge uniformemente a $\int_{\mathcal{X}} \mathcal{A}^Y(\Lambda_x f_x)(\cdot) dx$. Usando que A es un operador cerrado, se obtiene la ecuación (3.9).
Para el caso en el que $f \in C_0^\infty(\mathcal{X})$, la prueba procede igual haciendo las correspondientes correcciones notacionales, salvo en la ecuación (3.11), en donde se tiene que hacer uso de que Λ está acotada en el conjunto $\mathcal{Y} \times K$ para garantizar que las desigualdades se cumplen. □

Antes de enunciar el siguiente lema, se hará la siguiente observación: dada una sucesión de funciones $(f_k)_{k \in \mathcal{N}}$ en $C_0(\mathcal{X} \times \mathcal{Y})$ que converge uniformemente a una función f, se tiene

$$\int_{\mathcal{X}} \lambda(\cdot, x) f_k(x, \cdot) dx \to \int_{\mathcal{X}} \lambda(\cdot, x) f(x, \cdot) dx \text{ uniformemente en } \mathcal{Y}. \tag{3.12}$$

En efecto, esto se sigue al considerar el siguiente cálculo

$$\begin{aligned}\left| \int_{\mathcal{X}} \Lambda(y,x) f(x,y) dx - \int_{\mathcal{X}} \Lambda(y,x) f_n(x,y) dx \right| &= \left| \int_{\mathcal{X}} \Lambda(y,x)(f(x,y) - f_n(x,y)) dx \right| \\ &\leq \int_{\mathcal{X}} \Lambda(y,x) \|f - f_n\|_{\mathcal{X}\times\mathcal{Y}} dx \\ &= \|f - f_n\|_{\mathcal{X}\times\mathcal{Y}}\end{aligned}$$

y esta última expresión converge a 0 cuando n tiende a infinito. Un resultado similar se obtiene al fijar $x \in \mathcal{X}$ e integrar sobre $\mathcal{Y}$.

Lema 3.6. *Sea $f \in \mathcal{D}(\mathcal{A}^Z)$ y $g \in \mathcal{D}(\mathcal{A}^Y)$. Supóngase que para cierto $t_0 \geq 0$, la distribución condicional de $Z^1(t_0)$ dada $Z^2(t_0)$ es L. Entonces, para cada $t \geq 0$, las funciones*

$$\begin{aligned} u(t) : \mathcal{Y} \to \mathbb{R}, \quad & y \mapsto \mathbb{E}[f(Z^1(t+t_0), Z^2(t+t_0)) \mid Z^2(t_0) = y] \\ v(t) : \mathcal{Y} \to \mathbb{R}, \quad & y \mapsto \mathbb{E}[g(Z^2(t+t_0)) \mid Z^2(t_0) = y] \end{aligned} \tag{3.13}$$

están en el dominio de $\mathcal{A}^Y$, las funciones $t \mapsto u(t)$ y $t \mapsto v(t)$ son continuamente diferenciables con respecto a la norma uniforme en $C_0(\mathcal{Y})$ y

$$\frac{d}{dt} u(t) = \mathcal{A}^Y u(t) \quad \frac{d}{dt} v(t) = \mathcal{A}^Y v(t), \quad t \geq 0$$

En particular, bajo las hipótesis del Teorema 3.3, esto se cumple con $t_0 = 0$.

Demostración. Solo se probará el lema para u, pues la prueba para v es similar haciendo los justes necesarios.
Para cada $t \geq 0$, defínase la función

$$w(t) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}, \quad (x,y) \mapsto \mathbb{E}[f(Z^1(t+t_0), Z^2(t+t_0)) \mid Z^1(t_0) = x, Z^2(t_0) = y].$$

Como la distribución condicional de Z^1 dado Z^2 es Λ, se tiene entonces que

$$\begin{aligned} u(t)(Z^2(t_0)) &= \mathbb{E}[f(Z^1(t+t_0), Z^2(t+t_0)) \mid Z^2(t_0)] \\ &= \mathbb{E}[f(Z^1(t+t_0), Z^2(t+t_0)) \mid Z^1(t_0), Z^2(t_0) \mid Z^2(t_0)] \\ &= \mathbb{E}[w(Z^1(t_0), Z^2(t_0)) \mid Z^2(t_0)] \\ &= \int_{\mathcal{X}} w(x, Z^2(t_0)) \Lambda(Z^2(t_0), x) dx \end{aligned}$$

por lo que se puede tomar $u(t)(y) = \int_{\mathcal{X}} w(t)(x,y)\Lambda(y,x)dx$. Obsérvese que por homegeneidad del proceso Z, $w(t)$ es simplemente la acción del semigrupo de Z al tiempo t sobre f, por lo que el Teorema (2.4), implica que $w(t)$ pertenece al dominio de $\mathcal{A}^Z$ para cada $t \geq 0$ y se satisface

$$\dot{w}(t) = \mathcal{A}^Z w(t)$$

Ahora, por la observación hecha antes del presente lema, u también es diferenciable respecto a la norma uniforme en C_0 y su derivada está dada por

$$\dot{u}(t)(y) = \int_{\mathcal{X}} \Lambda(y,x)\dot{w}(t)(x,y)dx = \int_{\mathcal{X}} \Lambda(y,x)\mathcal{A}^Z w(t)(x,y)dx \tag{3.14}$$

Por la Condición 1, $C_0^\infty(\mathcal{X}\times\mathcal{Y})$ es un centro para $\mathcal{A}^Z$, por lo que existe una sucesión $(w_k(t))_{k\in\mathbb{N}}$ en $C_0^\infty(\mathcal{X}\times\mathcal{Y})$ que converge a $w(t)$ uniformemente en $\mathcal{X}\times\mathcal{Y}$ y tal que $\mathcal{A}^Z w_k(t)$ converge uniformemente a $\mathcal{A}^Z w(t)$. Por la definición de V se tiene que

$$\Lambda\mathcal{A}^Z = \Lambda\mathcal{A}^x + \Lambda\mathcal{A}^Y + \Lambda(\nabla_y V)'\rho\nabla_y = [\Lambda\mathcal{A}^Y + (\nabla_y\Lambda)'\rho\nabla_y + (\mathcal{A}^Y\Lambda_x)] + [\Lambda\mathcal{A}^x - (\mathcal{A}^Y\Lambda_x)]$$

Usando la ecuación (3.3) y los lemas 3.4 y 3.5, se sigue que la integral de la derecha en (3.14) se puede escribir como

$$\begin{aligned}
\int_{\mathcal{X}} \Lambda(y,x)\mathcal{A}^Z(w(t))_x(y)dx &= \lim_{k\to\infty}\int_{\mathcal{X}} \Lambda(y,x)\mathcal{A}^Z(w_k(t))_x(y)dx \\
&= \lim_{k\to\infty}\int_{\mathcal{X}} ([\Lambda\mathcal{A}^Y + (\nabla_y\Lambda)'\rho\nabla_y + (\mathcal{A}^Y\Lambda_x)] + [\Lambda\mathcal{A}^X - (\mathcal{A}^Y\Lambda_x)])(w_k(t))_x(y)dx \\
&= \lim_{k\to\infty}\int_{\mathcal{X}} [\Lambda\mathcal{A}^Y + (\nabla_y\Lambda)'\rho\nabla_y + (\mathcal{A}^Y\Lambda_x)](w_k(t))_x(y)dx \\
&= \lim_{k\to\infty}\int_{\mathcal{X}} \mathcal{A}^Y(\Lambda_x(w_k)(t)_x)(y)dx \\
&= \lim_{k\to\infty}\mathcal{A}^Y\int_{\mathcal{X}} \Lambda_x(w_k(t))_x(y)dx
\end{aligned} \tag{3.15}$$

donde el límite se entiende de manera uniforme en y. Usando entonces que A es cerrado, se sigue $u(t) = \int_{\mathcal{X}} \Lambda_x(w_k(t))_x(\cdot)dx$ pertenece a $\mathcal{D}(\mathcal{A}^Y)$ y

$$\mathcal{A}^Y u(t) = \int_{\mathcal{X}} \Lambda(y,x)\mathcal{A}^Z(w(t))_x(\cdot)dx = \dot{u}(t).$$

$\square$

Lema 3.7. *Z^1 es un proceso de Feller con respecto a la filtración $(\mathscr{F}_t^X)_{t\geq 0}$.*

Demostración. Sea $f \in C_0^\infty(\mathcal{X})$. Si $\pi_m : \mathcal{X}\times\mathcal{Y}\to\mathcal{X}$ es la proyección canónica en $\mathcal{X}$, la función $f\circ\pi_m$ pertenece a $C_0^\infty(\mathcal{X}\times\mathcal{Y})$, por lo que la fórmula de Itô implica que

$$\begin{aligned}
d(f(Z_s^1)) = d(f\circ\pi_m(Z_s)) &= \sum_{i=1}^{m+n}\frac{\partial(f\circ\pi_m)}{\partial z_i}(Z_s)dZ_s^i + \frac{1}{2}\sum_{i,j=1}^{m} a_{ij}(Z_s^1)\frac{\partial^2(f\circ\pi_m)}{\partial x_i\partial x_j}(Z_s)ds \\
&\quad + \frac{1}{2}\sum_{k,l=1}^{n}\rho_{kl}(Z_s^2)\frac{\partial^2(f\circ\pi_m)}{\partial y_k\partial y_l}(Z_s)ds.
\end{aligned}$$

Aquí, las primeras m parciales de $f \circ \pi_m$ coinciden con las parciales de f y últimas n parciales de $f \circ \pi_m$ se anulan, por lo que esta expresión se reduce a

$$d(f(Z_s^1)) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(Z_s^1) d(Z_s^1)^i + \frac{1}{2}\sum_{i,j=1}^{m} a_{ij}(Z_s^1)\frac{\partial^2 f}{\partial x_i \partial x_j}(Z_s^1)$$

Lo cual implica que Z^1 satisface el problema de la martingala asociado a $\mathcal{A}^X$. Como se tiene que $X(0)$ y Z_0^1 tienen la misma distribución, se debe tener entonces que Z^1 y X son idénticamente distirbuidos. En particular, como el problema asociado a $\mathcal{A}^X$ está bien definido, Z^1 es de Feller respecto a su filtración. $\square$

En los siguientes lemas, se hará uso de que cualquier función medible acotada puede ser aproximada por una sucesión de funciones en C_0^∞ unifomemente acotada. Usandoel Teorema de la convergencia dominada, solo hace falta ver que los enunciados se satisfacen cuando se restringe a este conjunto.

Lema 3.8. *Los semigrupos de transición $(Q_t)_{t\geq 0}$ y $(P_t)_{t\geq 0}$ satisfacen la relación $Q\langle L\rangle P$.*

Demostración. Solo hace falta ver que para cualquier función $f \in C_0^\infty(\mathcal{X})$ se satisface

$$Q_t L f = L P_t f$$

Usando la definición de L, esto es equivalente a mostrar que para cualquier $t \geq 0$, se cumple

$$\mathbb{E}\left[\int_{\mathcal{X}} \Lambda(Z_t^2, x) f(x) dx \,\middle|\, Z_0^2 = y\right] = \int_{\mathcal{X}} \Lambda(x,y)\mathbb{E}[f(Z_t^1) \mid Z_0^1 = x] dx \tag{3.16}$$

Con este fin, defínase

$$\begin{aligned} u(t) : \mathcal{Y} \to \mathbb{R}, &\quad y \mapsto \mathbb{E}\left[\int_{\mathcal{X}} \Lambda(Z_t^2, x) f(x) dx \,\middle|\, Z_0^2 = y\right] \\ v(t) : \mathcal{X} \to \mathbb{R}, &\quad x \mapsto \int_{\mathcal{X}} \Lambda(x,y)\mathbb{E}[f(Z_t^1) \mid Z_0^1 = x] dx \end{aligned}$$

Por el Lema 3.6, $u(0)$ (que está dada por $y \mapsto \int_{\mathcal{X}} \Lambda(x,y) f(x) dx$) pertenece al dominio de $\mathcal{A}^Y$. Por lo tanto, el otra aplicación del mismo lema implica que u es una solución al problema de Cauchy con valor inicial

$$\dot{u}(t) = \mathcal{A}^Y u(t), \quad u(0)(y) = \int_{\mathcal{X}} \Lambda(x,y) f(x) dx \tag{3.17}$$

Por el Teorema 2.15, la solución a (3.17) es única y para probar (3.16) solo hace falta ver entonces que el lado derecho de (3.16), dado por $L(v(t))$, también resuelve (3.17).

Por el Lema 3.7 y el Teorema 2.4, v es diferenciable con respecto a la norma uniforme en $C_0(\mathcal{X})$ y $\dot{v}(t) = \mathcal{A}^X v(t)$ para todo $t \geq 0$.
Ahora, si $(v_l(t))_{l\in\mathbb{N}}$ es una sucesión de funciones en $C_0^\infty(\mathcal{X})$ tal que $v_l(t) \to v(t)$ uniformemente en $\mathcal{X}$ y $\mathcal{A}^X v_l(t) \to \mathcal{A}^X v(t)$ también uniformemente, se tiene que la observación hecha antes del Lema 3.6 implica que

$$Lv_l(t) \to Lv(t) \text{ y } L(\mathcal{A}^X v_l(t)) \to L(\mathcal{A}^X v(t)) \text{ uniformemente en } \mathcal{Y}. \tag{3.18}$$

Sin embargo, (3.3) y el Lema 3.5 implica que, para cada $l \in \mathbb{N}$,

$$\begin{aligned} L(\mathcal{A}^X v_l(t))(y) = \int_{\mathcal{X}} \Lambda(y,x)(\mathcal{A}^X v_l(t))(x)dx &= \int_{\mathcal{X}} (\mathcal{A}^Y \Lambda_x)(y) v_l(t)(x)dx \\ &= \mathcal{A}^Y \int_{\mathcal{X}} \Lambda(x,y) v_l(t)(x)dx = \mathcal{A}^Y (Lv_l(t))(y) \end{aligned}$$

De donde (3.18) y el que $\mathcal{A}^Y$ sea cerrado implican que $L(\mathcal{A}^X v(t)) = \mathcal{A}^Y(Lv(t))$. Usando una vez más la observación antes del Lema 3.6, se obtiene

$$\frac{d}{dt} Lv(t) = L\dot{v}(t) = L(\mathcal{A}^X v(t)) = \mathcal{A}^Y(Lv(t))$$

lo que significa que Lv también satisface (3.17). □

Lema 3.9. *Para cualquier $t \geq 0$, condicionado a Z_t^2, la variable aleatoia Z_t^1 es independiente de $\mathscr{F}_t^Y$ y se distibuye condicionalmente de acuerdo a L.*

Demostración. Obsérvese que probar esto es equivalente a probar que para cualquier función $f \in C_0^\infty(\mathcal{X})$ y cualesquiera $0 = t_0 < t_1 < \cdots < t_k = t$, si $\mathcal{G} = \sigma(Z_0^2, Z^2(t_1), \ldots, Z^2(t_k))$, se satisface

$$\mathbb{E}[f(Z_t^1) \mid \mathcal{G}] = (Lf)(Z_t^2)$$

La prueba se hará por inducción. Para $k = 1$, se debe mostrar que si $g \in C_0^\infty(\mathcal{Y})$, entonces

$$\mathbb{E}[f(Z_t^1)g(Z_t^2) \mid Z_0^2] = \mathbb{E}[(Lf)(Z_t^2)g(Z_t^2) \mid Z_0^2] \tag{3.19}$$

Para dichas f y g, defínase

$$\begin{aligned} u(t) &: \mathcal{Y} \to \mathbb{R}, \quad y \mapsto \mathbb{E}[f(Z_t^1)g(Z_t^2) \mid Z_0^2 = y] \\ v(t) &: \mathcal{Y} \to \mathbb{R}, \quad y \mapsto \mathbb{E}[(Lf)(Z_t^2)g(Z_t^2) \mid Z_0^2 = y] \end{aligned}$$

Se desea mostrar que las funciones u y v son iguales y , por el Teorema 2.15, solo hace falta ver que ambas satisfacen el mismo problema de Cauchy con valor inicial.
Se tiene $u(0) = v(0)$, pues, por hipótesis, la distribución condicional de Z_0^1 dado Z_0^2 es Λ. Por

los lemas 3.5 y 3.4, la función $(Lf)g$ pertenece al dominio de $\mathcal{A}^Y$, por lo que el Lema 3.6 y los comentarios posteriores a su demostración implican que v satisface la ecuación

$$\dot{v}(t) = \mathcal{A}^Y v(t)$$

Defínase $h : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ dada por $h(x, y) = f(x)g(y)$. Así, $h \in C_0^\infty(\mathcal{X} \times \mathcal{Y}) \subset \mathcal{D}(\mathcal{A}^Z)$, por lo que otra aplicación del Lema 3.6 implica que se satisface

$$\dot{u}(t) = \mathcal{A}^Y u(t)$$

que era lo que se quería probar.
Supóngase ahora que (3.19) es válida para algún $k \in \mathbb{N}$. Así, la esperanza condicional de $Z^1(t_k)$ dado $(Z_0^2, \ldots, Z^2(t_k))$ es de nuevo L. Para probar el paso inductivo, se puede repetir el argumento para $k = 1$ con el proceso de Feller $(Z(t_k + t))_{t \geq 0}$ después de condicionar con $(Z_0^2, \ldots, Z^2(t_k))$, lo cual finaliza la prueba. $\square$

Lema 3.10. *Los procesos Z^2 y Y tienen la misma distribución.*

Demostración. Solo es necesario ver que se tienen las correctas funciones de transición, es decir, para cualquier $f \in C_0^\infty(\mathcal{Y})$ y $0 \leq s, t$ se tiene

$$\mathbb{E}[f(Z_{t+s}^2) \mid \mathscr{F}_s^Y] = (Q_t f)(Z_s^2) \tag{3.20}$$

Defínanse

$$\begin{aligned} u(t) : \mathcal{Y} \to \mathbb{R}, &\quad y \mapsto \mathbb{E}[f(Z_{t+s}^2) \mid Z_s^2 = y] \\ w(t) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}, &\quad (x, y) \mapsto \mathbb{E}[f(Z_{t+s}^2) \mid Z_s^1 = x, Z_s^2 = y]. \end{aligned}$$

Usando que Z es un proceso de Markov respecto a $(\mathscr{F}_t)_{t \geq 0}$ y que $f(Z_t^2)$ es función de Z_t, el Lema 3.9, implica lo siguiente

$$\begin{aligned} \mathbb{E}[f(Z_t^2) \mid \mathscr{F}_s^Y] &= \mathbb{E}[f(Z_t^2) \mid \mathscr{F}_s \mid \mathscr{F}_s^Y] \\ &= \mathbb{E}[f(Z_t^2) \mid Z_s^1, Z_s^2 \mid \mathscr{F}_s^Y] \\ &= \mathbb{E}[w(t)(Z_s^1, Z_s^2) \mid \mathscr{F}_s^Y] \\ &= \mathbb{E}[w(t)(Z_s^1, Z_s^2) \mid Z_s^2] = \int_{\mathcal{X}} w(x, Z_s^2) \Lambda(Z_s^2, x) dx \end{aligned}$$

Sin embargo, de la demostación del Lema 3.6, se ve que también se tiene la igualdad

$$u(t)(Z_s^2) = \int_{\mathcal{X}} w(x, Z_s^2) \Lambda(Z_s^2, x) dx$$

por lo que $u(t)(Z_s^2) = \mathbb{E}[f(Z_{t+s}^2) \mid \mathscr{F}_s^Y]$.
Por el Lema 3.6, u resuelve la ecuación

$$\dot{u}(t) = \mathcal{A}^Y u(t), \quad u(0) = f$$

y si se define v por $v(t) = Q_t f$, el Teorema 2.4 implica entonces que v también es diferenciable y satisface

$$\dot{v}(t) = \mathcal{A}^Y v(t), \quad v(0) = f$$

Por la unicidad enunciada en el Teorema 2.15, se sigue que $u(t) = v(t)$ para cada $t \geq 0$ y, en particular,

$$\mathbb{E}[f(Z_{t+s}^2) \mid \mathscr{F}_s^Y] = u(t)(Z_s^2) = v(t)(Z_s^2) = (Q_t f)(Z_s^2)$$

como se quería. $\square$

Lema 3.11. *El proceso Z^1 es un proceso de Markov con respecto a la filtración generada por Z.*

Demostración. Fíjese $0 \leq s < t$. Se debe demostrar que, para cualquier función $f \in C_0^\infty(\mathcal{X})$, se tiene

$$\mathbb{E}[f(Z_t^1) \mid \mathscr{F}_s^Z] = \mathbb{E}[f(Z_t^1) \mid Z_s^1]$$

Como Z es un proceso homogéneo de Markov y $f(Z_t^1)$ es función de Z_t, se tiene

$$\mathbb{E}[f(Z_t^1) \mid \mathscr{F}_s^Z] = \mathbb{E}[f(Z_t^1) \mid Z_s^1, Z_s^2] = \mathbb{E}[f(Z^1(t-s)) \mid Z_0^1, Z_0^2]$$

Por lo que solo es necesario demostrar que, condicionado a Z_s^1, Z_t^1 es independiente de Z_s^2 y por homegeneidad, se puede considerar $s = 0$.
Defínase

$$\begin{aligned} v(t) &: \mathcal{X} \to \mathbb{R}, \quad x \mapsto \mathbb{E}[f(Z_t^1) \mid Z_0^1 = x] \\ w(t) &: \mathcal{X} \times \mathcal{Y} \to \mathbb{R}, \quad (x, y) \mapsto \mathbb{E}[f(Z_t^1) \mid Z_0^1 = x, Z_0^2 = y]. \end{aligned}$$

Usando la observarción previa al Lema 3.7, se puede considerar a f como una función en $\mathcal{D}(\mathcal{A}^Z)$, por lo que el Teorema 2.4 implica que se satisface

$$\dot{u}(t) = \mathcal{A}^X u(t), \quad \dot{w}(t) = \mathcal{A}^Z w(t), \quad u(0) = w(0) = f$$

Y aún otra aplicación de la misma observación implica que se tiene $\mathcal{A}^X u(t) = \mathcal{A}^Z u(t)$, por lo que el Teorema 2.15 implica entonces que $u(t) = w(t)$, que a su vez implica lo que se quiere demostrar. $\square$

Por el punto (iii) en la Definición 3.2, se tiene, para f y g funciones medibles acotadas en $\mathcal{X}$ y $t \geq 0$,

$$\begin{aligned} \mathbb{E}[f(Z_0^1) g(Z_t^1) \mid Z_0^2] &= \mathbb{E}[f(Z_0^1) \mathbb{E}[g(Z_t^1) \mid Z_0^1, Z_0^2] \mid Z_0^2] \\ &= \mathbb{E}[f(Z_0^1)(P_t g)(Z_0^1) \mid Z_0^2] \\ &= L(f(P_t g))(Z_0^2) \end{aligned}$$

Cuando $f = 1_A$ y $g = 1_B$ para ciertos conjuntos $A, B \in \mathscr{B}(\mathcal{X})$, esto se puedo escribir como

$$P[Z_0^1 \in A, Z_t^1 \in B \mid Z_0^2] = \int_A P_t(x, B)\Lambda(Z_0^2, x)dx \tag{3.21}$$

lo que da una expresión para la distribución condicional de (Z_0^1, Z_t^1). Similarmente, por la condición (iv) de la Definición 3.2, se tiene, para g y h funciones medibles y acotadas en $\mathcal{X}$ y $\mathcal{Y}$ respectivamente,

$$\begin{aligned}\mathbb{E}[g(Z_t^1)h(Z_t^2) \mid Z_0^2] &= \mathbb{E}[h(Z_t^2)\mathbb{E}[g(Z_t^1) \mid \sigma(Z_s^2, 0 \leq s \leq t)] \mid Z_0^2] \\ &= \mathbb{E}[h(Z_t^2)(Lg)(Z_t^2) \mid Z_0^2] \\ &= Q_t(h(Lg))(Z_0^2)\end{aligned}$$

por lo que al tomar $g = 1_B$ y $h = 1_C$ para ciertos $B \in \mathscr{B}(\mathcal{X})$ y $C \in \mathscr{B}(\mathcal{Y})$, se puede escribir

$$P[Z_t^1 \in B, Z_t^2 \in C \mid Z_0^2] = \int_C L(y, B)Q_t(Z_0^2, dy) = \int_B \int_C \Lambda(y, x)Q_t(Z_0^2, dy)\, dx \tag{3.22}$$

por lo que en particular Z^1 tiene distribución condicional dada por $\int_B \int_{\mathcal{Y}} \Lambda(y, x)Q_t(Z_0^2, dy)\, dx$. De esto vemos que la medida $P[Z_t^1 \in \cdot, Z_t^2 \in C \mid Z_0^2]$ es absolutamente continua respecto a la medida $P[Z_t^1 \in \cdot \mid Z_0^2]$ y esta última es absolutamente continua respecto a la medida de Lebesgue en $\mathcal{X}$. Así, la derivada de Radon-Nikodym de $P[Z_t^1 \in \cdot, Z_t^2 \in C \mid Z_0^2]$ respecto a $P[Z_t^1 \in \cdot \mid Z_0^2]$ está dada por el cociente de las derivadas respecto a la medida de Lebesgue, es decir, por

$$\nu(Z_0^2, x, C) = \frac{\int_C \Lambda(y, x)Q_t(Z_0^2, dy)}{\int_{\mathcal{Y}} \Lambda(\tilde{y}, x)Q_t(Z_0^2, d\tilde{y})} = \int_C \frac{\Lambda(y, x)Q_t(Z_0^2, dy)}{\int_{\mathcal{Y}} \Lambda(\tilde{y}, x)Q_t(Z_0^2, d\tilde{y})} \tag{3.23}$$

Obsérverse que por definición de ν, se satisface la relación $\nu(Z_0^2, Z_t^1, C) = P[Z_t^2 \in C \mid Z_t^1, Z_0^2]$. Ahora, si dadas Z_t^1 y Z_0^2, Z_0^1 y Z_t^2 son independientes, entonces, para $A, B \in \mathscr{B}(\mathcal{X})$ y $C \in \mathscr{B}(\mathcal{Y})$, las ecuaciones (3.21) y (3.23) implican que, por desintegración,

$$\begin{aligned}P[Z_0^1 \in A, Z_t^1 \in B, Z_t^2 \in C \mid Z_0^2] &= \mathbb{E}[1_A(Z_0^1)1_B(Z_t^1)\mathbb{E}[1_C(Z_t^2) \mid Z_0^1, Z_t^1, Z_0^2] \mid Z_0^2] \\ &= \mathbb{E}[1_A(Z_0^1)1_B(Z_t^1)\mathbb{E}[1_C(Z_t^2) \mid Z_t^1, Z_0^2] \mid Z_0^2] \\ &= \int_A \int_B \nu(Z_0^2, \tilde{x}, C)P_t(x, d\tilde{x})\Lambda(Z_0^2, x)dx \\ &= \int_A \int_B \int_C \frac{\Lambda(y_1, x_0)Q_t(Z_0^2, dy_1)}{\int_{\mathcal{Y}} \Lambda(y_0, x_0)Q_t(Z_0^2, dy_0)} P_t(x, dx_0)\Lambda(Z_0^2, x)dx\end{aligned}$$

Definiendo a $(R_t)_{t\geq 0}$ como el semigrupo de transición de Z, se satisface

$$R_t(Z_0^1, Z_0^2, B \times C) = P[Z_t^1 \in B, Z_t^2 \in C \mid Z_0^1, Z_0^2],$$

por lo que, al ser L la distribución condicional de Z_0^1 dado Z_0^2, se sigue

$$\begin{aligned} P[Z_0^1 \in A, Z_t^1 \in B, Z_t^2 \in C \mid Z_0^2] &= \mathbb{E}[1_A(Z_0^1)\mathbb{E}[1_B(Z_t^1)1_C(Z_t^2) \mid Z_0^1, Z_0^2] \mid Z_0^2] \\ &= \int_A R_t(x, Z_0^2, B \times C) L(Z_0^2, dx) \end{aligned}$$

De donde concluimos que la densidad de $P[Z_0^1 \in \cdot, Z_t^1 \in B, Z_t^2 \in C \mid Z_0^2]$ respecto a $L(Z_0^2, \cdot)$ es $R_t(\cdot, Z_0^2, B \times C)$. Pero estas dos medidas son de absolutamente continuas respecto a a la medida de Lebesgue, por lo que debe tenerse

$$\begin{aligned} R_t(x, Z_0^2, B \times C) &= \frac{\Lambda(Z_0^2, x) \int_B \int_C \frac{\Lambda(y_1,x_1)Q_t(Z_0^2,dy_1)}{(Q_t\Lambda_{x_1})(Z_0^2)} P_t(Z_0^1, dx_1)}{\Lambda(Z_0^2, x)} \\ &= \int_B \int_C \frac{\Lambda(y_1, x_1)Q_t(Z_0^2, dy_1)}{\int_{\mathcal{Y}} \Lambda(\tilde{y}_1, x_1)Q_t(Z_0^2, d\tilde{y}_1)} P_t(Z_0^1, dx_1) \end{aligned} \tag{3.24}$$

En conclusión, si el punto (v) de la Definición 3.2 se satisface, entonces la el semigrupo de transición tiene la forma expresada en la ecuación (3.24) y, siguiendo los pasos en orden inverso, se tiene que, de hecho, la ecuación (3.24) implica el punto (v) en la Definición 3.2, por lo que son afirmaciones equivalentes. Considerando que los puntos (i) a (iv) ya han sido probados en los lemas anteriores, para finalizar la prueba del Teorema 3.3, solo hace falta probar la ecuación (3.24). Algunos otros lemas serán necesarios para este fin.

Lema 3.12. *Al tiempo $t \geq 0$, la ley de la difusión Z es absolutamente continua con respecto a la ley de la difusión producto $\widehat{Z} = (\widehat{Z}^1, \widehat{Z}^2)$ en $\mathcal{X} \times \mathcal{Y}$ con generador $\mathcal{A}^X + \mathcal{A}^Y$ y condición inicial $\widehat{Z}_0 = Z_0$ y la densidad correspondiente está dada por*

$$\exp\left(\int_0^t \left((\nabla_y V)(\widehat{Z}_s^2, \widehat{Z}_s^1)\right)' d\widehat{M}_2(s) - \frac{1}{2} \int_0^t \|(\nabla_y V)(\widehat{Z}_s^2, \widehat{Z}_s^1)\|^2_{\rho(\widehat{Z}_s^2)} ds \right) \tag{3.25}$$

donde $\widehat{M_2}$ es la parte de martingala local de $\widehat{Z}^2$ y $\|y\|^2_{\rho(\widehat{Z}_s^2)} = y'\rho(\widehat{Z}_s^2)y$.

Demostración. En el espacio filtrado, $(C([0,\infty), \mathcal{X} \times \mathcal{Y}), \mathscr{F}, (\mathscr{F}_t)_{t\geq 0})$ considérese la distribución de $\widehat{Z}$, $\widehat{P}$. Si $\sigma : \mathcal{Y} \to \mathbb{R}^{n\times n}$ es tal que $\sigma\sigma' = \rho$, entonces $\widehat{Z}^2$ satisface la ecuación

$$\widehat{Z}_t^2 = \widehat{Z}_0^2 = \int_0^t \sigma(\widehat{Z}_s^2) dB_s + \int_0^t \gamma(\widehat{Z}_s^2) ds$$

donde $(B_t)_{t\geq 0} = (B_t^1, B_t^2, \ldots, B_t^n)_{t\geq 0}$ es un movimiento Browniano n-dimensional respecto a $\widehat{P}$. Así, de la definición de $\widehat{M_2}$ se sigue entonces que $\widehat{M_2} = \widehat{Z}_0^2 + \int_0^\cdot \sigma(\widehat{Z}_s^2) dB_s$.
Para simplificar la notación, sea $K(s) = (\nabla_y V)(\widehat{Z}_s^2, \widehat{Z}_s^1)$. Defínanse los tiempos de paro

$$\tau_n = \text{ínf}\left\{ t \geq 0 \;\middle|\; \int_0^t \|K(s)\|^2_{\rho(\widehat{Z}_s^2)} ds \geq n \right\}, \quad n \in \mathbb{N}$$

y la martingala local L dada por

$$L_t = K(s) \cdot \widehat{M}_2 = \sum_{i=1}^{n} K^i \cdot \widehat{M}_2^i$$

donde K^i es la i-ésima entrada del proceso K y similarmente se define $\widehat{M}_2^i$. Así,

$$\langle \widehat{M}_2^i, \widehat{M}_2^j \rangle = \sum_{k=1}^{n} \sum_{l=1}^{n} \sigma_{ik}\sigma_{jl} \cdot \langle B^k, B^l \rangle = \sum_{k=1}^{n} \sigma_{ik}\sigma_{jk} = \rho_{ij}(\widehat{Z}^2)$$

Por lo que

$$\langle L, L \rangle_t = \sum_{i=1}^{n} \sum_{k=1}^{n} K^i K^j \cdot \langle \widehat{M}_2^i, \widehat{M}_2^j \rangle = \sum_{i=1}^{n} \sum_{k=1}^{n} K^i K^j \cdot \rho_{ij}(\widehat{Z}_t^2) = \int_0^t \|K(s)\|^2_{\rho(\widehat{Z}_s^2)} ds$$

De la definición de los tiempos de paro, se concluye que $\langle L^{\tau_n}, L^{\tau_n} \rangle_\infty = \langle L, L \rangle_{\tau_n} \leq n$. Por lo tanto

$$\mathbb{E}\left[\exp\left(\frac{1}{2} \langle L^{\tau_n}, L^{\tau_n} \rangle_\infty \right) \right] \leq \exp\left(\frac{n}{2} \right) < \infty$$

El criterio de Novikov implica que el proceso $D^{(n)}$ definido por

$$\begin{aligned} D_t^{(n)} &= \exp\left(L_t^{\tau_n} - \frac{1}{2} \langle L^{\tau_n}, L^{\tau_n} \rangle_t \right) \\ &= \exp\left(\int_0^{t\wedge\tau_n} K(s) d\widehat{M}_2(s) - \frac{1}{2} \int_0^{t\wedge\tau_n} \|K(s)\|^2_{\rho(\widehat{Z}_s^2)} ds \right) \end{aligned}$$

es una martingala uniformemente integrable y se puede definir una medida $Q^{(n)}$ en $\mathscr{F}$ de manera que las restricciones a las σ-álgebras $\mathscr{F}_t$ satisfacen $Q^{(n)} = D_t^{(n)} \cdot \widehat{P}$ y, como una consecuencia del Teorema de Girsanov y del Teorema de caracterización de Levy, el proceso $\widehat{B} = (\widehat{B}^1, \ldots, \widehat{B}^n)$ definido por $\widehat{B}^i = B^i - \langle B^i, L \rangle$ es un movimiento Browniano respecto a $Q^{(n)}$.
Como $\langle B^j, L \rangle = \int_0^\cdot \sum_{k=1}^n K^k(s)\sigma_{kj}(\widehat{Z}_s^2) ds$, entonces, para cada $i \in \{1, \ldots, n\}$,

$$\begin{aligned} \sum_{j=1}^{n} \int_0^t \sigma_{ij}(\widehat{Z}_s^2) dB_s^j &= \sum_{j=1}^{n} \int_0^t \sigma_{ij}(\widehat{Z}_s^2) d\widehat{B}_s^j + \int_0^t \sum_{j=1}^{n} \sigma_{ij}(\widehat{Z}_s^2) d\langle B^j, L \rangle \\ &= \sum_{j=1}^{n} \int_0^t \sigma_{ij}(\widehat{Z}_s^2) d\widehat{B}_s^j + \int_0^t \sum_{j=1}^{n} \sum_{k=1}^{n} K^k(s) \sigma_{ij}(\widehat{Z}_s^2) \sigma_{kj}(\widehat{Z}_s^2) d\langle B^j, L \rangle \\ &= \sum_{j=1}^{n} \int_0^t \sigma_{ij}(\widehat{Z}_s^2) d\widehat{B}_s^j + \int_0^t \sum_{k=1}^{n} K^k(s) \rho_{ik}(\widehat{Z}_s^2) ds \end{aligned}$$

Escrito en forma diferencial, estas igualdades dicen que

$$\sigma(\widehat{Z}_t^2)dB_t = \sigma(\widehat{Z}_t^2)d\widehat{B}_t + K(s)\rho(\widehat{Z}_t^2)dt$$

Y recordando la definición de K y de $\widehat{Z}$, esto significa que $\widehat{Z}$ satisface la ecuación diferencial estocástica para Z en el intervalo $[0, t \wedge \tau_n]$. Por la Condición 1, la ley de dicha solución es única, por lo que en los conjuntos $\{\tau_n \geq t\}$, $P = \widehat{Z}(Q^{(n)}) = Q^{(n)} \ll \widehat{P}$ (recuérdese que $\widehat{Z}$ es el proceso canónico) y la densidad está dada por (3.25).
Por la continuidad de $\nabla_y V$ y ρ, con probabilidad uno se tiene

$$\int_0^t \|K(s)\|^2_{\rho(\widehat{Z}_s^2)} ds < \infty$$

por lo que los eventos $\{\tau_n \geq 1\}$ crecen a un evento con probabilidad 1 respecto a $\widehat{P}$, lo cual concluye la demostración. □

Lema 3.13. *En la notación del lema anterior, para cada $t \geq 0$, las densidades definidas en* (3.25) *se pueden escribir como*

$$\lim_{\varepsilon \downarrow 0} \prod_{j=1}^{J(\varepsilon)} \frac{\Lambda(\widehat{Z}^2_{t_j^\varepsilon}, \widehat{Z}^1_{t_j^\varepsilon})}{(Q_{t_j^\varepsilon - t_{j-1}^\varepsilon} \Lambda_{\widehat{Z}^1_{t_j^\varepsilon}})(\widehat{Z}^2_{t_{j-1}^\varepsilon})} \tag{3.26}$$

a lo largo de una subsucesión adecuada, donde, para $\varepsilon > 0$, $J(\varepsilon)$ es el menor entero mayor o igual que t/ε y $t_j^\varepsilon = (j\varepsilon) \wedge t$ para $j = 1, 2, \ldots, J(\varepsilon)$.

Demostración. Primero, se afirma que las densidades en (3.25) se pueden escribir como el límite casi seguro

$$\lim_{\varepsilon \downarrow 0} \exp\left(\sum_{j=1}^{J(\varepsilon)} \int_{t_{j-1}^\varepsilon}^{t_j^\varepsilon} (\nabla_y V)(\widehat{Z}_s^2, \widehat{Z}^1_{t_k^\varepsilon}) d\widehat{M}_2(s) - \frac{1}{2} \sum_{j=1}^{J(\varepsilon)} \int_{t_{j-1}^\varepsilon}^{t_j^\varepsilon} \|(\nabla_y V)(\widehat{Z}_s^2, \widehat{Z}^1_{t_k^\varepsilon})\|^2_{\rho(\widehat{Z}_s^2)} ds \right) \tag{3.27}$$

a lo largo de una subsucesión adecuada. En efecto, por el Teorema 2.5, la diferencia entre las integrales estocásticas en (3.27) y las análogas en (3.25) se puede escribir como un movimiento browniano estándar evaluado en la variación cuadrática de esa diferencia. Calculos similares a los hechos en el lema anterior dan una expresión exacta para estas variaciones cuadráticas, por lo que, gracias a la continuidad $\nabla_y V$ y ρ en el intervalo $[0, t]$, todas las variaciones cuadráticas convergen simultaneamente a cero casi seguramente cuando ε tiendo a 0. Se sigue que la diferencia entre las integrales converge a cero en probabilidad y, por lo tanto, a lo largo de alguna subsucesión. La continuidad uniforme de $\nabla_y V$ y el acotamiento uniforme de ρ en conjuntos compactos muestra también que la diferencia entre los términos de variación acotada en los exponentes en (3.25) y en

(3.27) tiende a cero casi seguramente, lo que da la representación en (3.27).
Ahora, para cada $j = 1, 2, \dots, J(\varepsilon)$, se tiene

$$\begin{aligned} \log \Lambda(\widehat{Z}^2_{t^\varepsilon_j}, \widehat{Z}^1_{t^\varepsilon_j}) - \log \Lambda(\widehat{Z}^2_{t^\varepsilon_{j-1}}, \widehat{Z}^1_{t^\varepsilon_j}) &= \int_{t^\varepsilon_{j-1}}^{t^\varepsilon_j} (\nabla_y V)(\widehat{Z}^2_s, \widehat{Z}^1_{t^\varepsilon_j}) d\widehat{M}_2(s) \\ &+ \int_{t^\varepsilon_{j-1}}^{t^\varepsilon_j} \frac{\mathcal{A}^Y \Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}}{\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}} (\widehat{Z}^2_s) ds - \frac{1}{2} \int_{t^\varepsilon_{j-1}}^{t^\varepsilon_j} \|(\nabla_y V)(\widehat{Z}^2_s, \widehat{Z}^1_s)\|^2_{\rho(\widehat{Z}^2_s)} ds \end{aligned} \tag{3.28}$$

Para ver esto, considérese, como en el Lema 3.4, una aproximación de $\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}$ por una sucesión de funciones en $C_0(\mathcal{Y})$ infinitamente diferenciables $(\varphi_{q, \widehat{Z}^1_{t^\varepsilon_j}})_{q \in \mathbb{N}}$ tales que

$$\varphi_{q, \widehat{Z}^1_{t^\varepsilon_j}} \to \Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}, \quad \nabla \varphi_{q, \widehat{Z}^1_{t^\varepsilon_j}} \to \nabla \Lambda_{\widehat{Z}^1_{t^\varepsilon_j}} \quad \text{y} \quad \mathcal{A}^Y \varphi_{q, \widehat{Z}^1_{t^\varepsilon_j}} \to \mathcal{A}^Y \Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}.$$

donde todas las convergencias son uniformes en $\mathcal{Y}$. Como el generador de $\widehat{Z}^2$ es $\mathcal{A}^Y$, la regla de Itô implica que, para cada $q \in \mathbb{N}$,

$$\begin{aligned} \varphi_{q, \widehat{Z}^1_{t^\varepsilon_j}}(\widehat{Z}^2_{t^\varepsilon_j}) - \varphi_{q, \widehat{Z}^1_{t^\varepsilon_j}}(\widehat{Z}^2_{t^\varepsilon_{j-1}}) &= \int_{t^\varepsilon_{j-1}}^{t^\varepsilon_j} (\nabla \varphi_{q, \widehat{Z}^1_{t^\varepsilon_j}})(\widehat{Z}^2_s) d\widehat{M}_2(s) \\ &+ \int_{t^\varepsilon_{j-1}}^{t^\varepsilon_j} (\mathcal{A}^Y \varphi_{q, \widehat{Z}^1_{t^\varepsilon_j}})(\widehat{Z}^2_s) ds \end{aligned}$$

Por lo que al tomar el límite se obtiene

$$\begin{aligned} \Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}(\widehat{Z}^2_{t^\varepsilon_j}) - \Lambda_{\widehat{Z}^1(t^\varepsilon_j)}(\widehat{Z}^2_{t^\varepsilon_{j-1}}) &= \int_{t^\varepsilon_{j-1}}^{t^\varepsilon_j} (\nabla \Lambda_{\widehat{Z}^1_{t^\varepsilon_j}})(\widehat{Z}^2_s) d\widehat{M}_2(s) \\ &+ \int_{t^\varepsilon_{j-1}}^{t^\varepsilon_j} (\mathcal{A}^Y \Lambda_{\widehat{Z}^1_{t^\varepsilon_j}})(\widehat{Z}^2_s) ds \end{aligned}$$

Usando que Λ es estrictamente positivo, una nueva aplicación de la regla de Itô a esta igualdad implica (3.28). Sustituyendo (3.28) en (3.27) se obtiene

$$\lim_{\varepsilon \downarrow 0} \prod_{j=1}^{J(\varepsilon)} \frac{\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}(\widehat{Z}^2_{t^\varepsilon_j})}{\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}(\widehat{Z}^2_{t^\varepsilon_{j-1}})} \exp\left(- \int_{t^\varepsilon_{j-1}}^{t^\varepsilon_j} \frac{\mathcal{A}^Y \Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}}{\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}} (\widehat{Z}^2_s) ds \right). \tag{3.29}$$

Si $o(\varepsilon^n)$ se usa para denotar funciones tales que $\lim_{\varepsilon\to 0}\varepsilon^{-n}o(\varepsilon^n) = 0$, entonces, para cada $j = 1, 2, \ldots, J(\varepsilon)$, se tiene

$$\begin{aligned}
&\frac{1}{\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}(\widehat{Z}^2_{t^\varepsilon_{j-1}})}\exp\Bigg(-\int_{t^\varepsilon_{j-1}}^{t^\varepsilon_j}\frac{\mathcal{A}^Y\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}}{\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}}(\widehat{Z}^2_s)ds\Bigg)\\
&\qquad = \frac{1}{\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}(\widehat{Z}^2_{t^\varepsilon_{j-1}})\Big(1+\int_{t^\varepsilon_{j-1}}^{t^\varepsilon_j}\frac{\mathcal{A}^Y\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}}{\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}}(\widehat{Z}^2_s)ds+o(\varepsilon)\Big)}\\
&\qquad = \frac{1}{\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}(\widehat{Z}^2_{t^\varepsilon_{j-1}})+(\mathcal{A}^Y\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}})(\widehat{Z}^2_{t^\varepsilon_{j-1}})(t^\varepsilon_j-t^\varepsilon_{j-1})+o(\varepsilon)}\\
&\qquad = \frac{1}{\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}}(\widehat{Z}^2_{t^\varepsilon_{j-1}})+\int_0^{t^\varepsilon_j-t^\varepsilon_{j-1}}(Q_s\mathcal{A}^Y\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}})(\widehat{Z}^2_{t^\varepsilon_{j-1}})ds+o(\varepsilon)}\\
&\qquad = \frac{1}{(Q_{t^\varepsilon_j-t^\varepsilon_{j-1}}\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}})(\widehat{Z}^2_{t^\varepsilon_{j-1}})+o(\varepsilon)}
\end{aligned}$$

donde en la tercer igualdad se ha usado la propiedad (iii) en la Definición 2.2 y en la última igualdad se ha usado que, según el Teorema 2.4, $\frac{d}{dt}Q_t f = Q_t\mathcal{A}^Y f$ para cada $f \in C_0(\mathcal{Y})$ y $t \geq 0$.
Por la continuidad uniforme de $\mathcal{A}^Y\Lambda_x$ y de Λ en conjuntos compactos, y puesto que se asume que $\mathcal{A}^Y\Lambda_x$ es acotada en conjuntos de la forma $\mathcal{Y}\times K$ para cada $x \in \mathcal{X}$ y $K \subset \mathcal{X}$ compacto, todos los términos $o(\varepsilon)$ convergen a cero uniformemente en j (aunque, por supuesto, pueden depender en la trayectoria de $\widehat{Z}$). Luego, (3.28) se puede escribir como

$$\lim_{\varepsilon\downarrow 0}\prod_{j=1}^{J(\varepsilon)}\frac{\Lambda(\widehat{Z}^2(t^\varepsilon_j),\widehat{Z}^1_{t^\varepsilon_j})}{(Q_{t^\varepsilon_j-t^\varepsilon_{j-1}}\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}})(\widehat{Z}^2_{t^\varepsilon_{j-1}})+o(\varepsilon)}$$

Por continuidad y debido a que Λ es estrictamente positivo, existe un $\delta > 0$ (que puede depender de la trayectoria de $\widehat{Z}$) tal que $\delta < \Lambda(\widehat{Z}^2(s_2), \widehat{Z}^1(s_1))$ para cada $s_1, s_2 \in [0, t]$, lo cual implica que $\delta < (Q_{t^\varepsilon_j-t^\varepsilon_{j-1}}\Lambda_{\widehat{Z}^1_{t^\varepsilon_j}})(\widehat{Z}^2_{t^\varepsilon_{j-1}})$ para cada $j = 1, 2, \ldots, J(\varepsilon)$. Tomando el logaritmo del cociente entre este límite y (3.26), se obtiene

$$
\begin{aligned}
0 \leq \lim_{\varepsilon \downarrow 0} \log \left(\prod_{j=1}^{J(\varepsilon)} \left(1 + \frac{o(\varepsilon)}{(Q_{t_j^\varepsilon - t_{j-1}^\varepsilon} \Lambda_{\widehat{Z}^1_{t_j^\varepsilon}})(\widehat{Z}^2_{t_{j-1}^\varepsilon})} \right) \right) &= \lim_{\varepsilon \downarrow 0} \sum_{j=1}^{J(\varepsilon)} \log \left(1 + \frac{o(\varepsilon)}{(Q_{t_j^\varepsilon - t_{j-1}^\varepsilon} \Lambda_{\widehat{Z}^1_{t_j^\varepsilon}})(\widehat{Z}^2_{t_{j-1}^\varepsilon})} \right) \\
&\leq \lim_{\varepsilon \downarrow 0} \sum_{j=1}^{J(\varepsilon)} \log \left(1 + \frac{o(\varepsilon)}{\delta} \right) \\
&= \lim_{\varepsilon \downarrow 0} J(\varepsilon) o(\varepsilon) = \lim_{\varepsilon \downarrow 0} o(1) = 0
\end{aligned}
$$

de donde se sigue el lema. □

Demostración del Teorema 3.3. Como se mencionó anteriormente, el punto (i) de la Definición 3.2 ya han sido demostrados en los lemas 3.7 y 3.10, y los puntos (ii), (iii) y (iv) en los lemas 3.8, 3.11 y 3.9 respectivamente, por lo que solo hace falta probar el último punto, el cual es equivalente a la igualdad en la ecuación (3.24). Sea $(R_t)_{t \geq 0}$ los operadores definidos por el lado derecho de (3.24), es decir,

$$
R_t(x, y, C) = \int_C \frac{\Lambda(y_1, x_1)}{\int_{\mathcal{Y}} \Lambda(\tilde{y}_1, x_1) Q_t(y, d\tilde{y}_1)} Q_t(y, dy_1) \otimes P_t(x, dx_1)
$$

de manera que cada R_t es un núcleo de transición y, de hecho, la colección $(\tilde{R}_t)_{t \geq 0}$ forma un semigrupo de transición (informalmente, cada operador representa la probabilidad de ir de un estado del proceso Z a otro, de manera que primero se considera la componente x, se muestrea conforme a P_t y luego, utilizando la regla de Bayes, se muestrea la componente y condicionada al valor nuevo de la componente x y el antiguo de la componente y).
Sea $\mathcal{Z}$ un boreliano en $\mathcal{X} \times \mathcal{Y}$. Usando la notación del Lema 3.12, denotamos por $\widehat{\mathbb{E}}$ a la esperanza respecto a $\widehat{P}$. Así, como $\widehat{Z}_0 = Z_0$ y el semigrupo de transición de $\widehat{Z}$ es $(\widehat{R}_t)_{t \geq 0} = (P_t \otimes Q_t)_{t \geq 0}$, el

Lema 3.13 implica lo siguiente:

$$\begin{aligned}
P[Z_t \in \mathcal{Z} \mid Z_0] &= \widehat{\mathbb{E}}\left[\lim_{\varepsilon\downarrow 0} \prod_{j=1}^{J(\varepsilon)} \frac{\Lambda(\widehat{Z}^2_{t^\varepsilon_j}, \widehat{Z}^1_{t^\varepsilon_j})}{(Q_{t^\varepsilon_j - t^\varepsilon_{j-1}} \Lambda_{\widehat{Z}^1_{t^\varepsilon_j}})(\widehat{Z}^2_{t^\varepsilon_{j-1}})} 1_{\mathcal{Z}}(\widehat{Z}_t) \mid \widehat{Z}_0 \right] \\
&\le \liminf_{\varepsilon\downarrow 0} \widehat{\mathbb{E}}\left[\prod_{j=1}^{J(\varepsilon)} \frac{\Lambda(\widehat{Z}^2_{t^\varepsilon_j}, \widehat{Z}^1_{t^\varepsilon_j})}{(Q_{t^\varepsilon_j - t^\varepsilon_{j-1}} \Lambda_{\widehat{Z}^1_{t^\varepsilon_j}})(\widehat{Z}^2_{t^\varepsilon_{j-1}})} 1_{\mathcal{Z}}(\widehat{Z}_t) \mid \widehat{Z}_0 \right] \\
&= \liminf_{\varepsilon\downarrow 0} \int_{\mathcal{X}\times\mathcal{Y}} \frac{\Lambda(y_1, x_1)}{(Q_{t^\varepsilon_1}\Lambda_{x_1})(Z^2_0)} \widehat{R}_{t^\varepsilon_1}(Z_0, dZ^1) \cdots \\
&\qquad \cdots \int_{\mathcal{X}\times\mathcal{Y}} \frac{\Lambda(y_2, x_2)}{(Q_{t^\varepsilon_2 - t^\varepsilon_1}\Lambda_{x_2})(y_1)} \widehat{R}_{t^\varepsilon_2 - t^\varepsilon_1}(Z^1, dZ^2) \cdots \\
&\qquad\quad \cdots \int_{\mathcal{X}\times\mathcal{Y}} \frac{1_{\mathcal{Z}}(z_{J(\varepsilon)})\Lambda(y_{J(\varepsilon)}, x_{J(\varepsilon)})}{Q_{t - t^\varepsilon_{J(\varepsilon)-1}} \Lambda_{x_{J(\varepsilon)}}(y_{J(\varepsilon)-1})} \widehat{R}_{t - t^\varepsilon_{J(\varepsilon)-1}}(z_{J(\varepsilon)-1}, dz_{J(\varepsilon)}) \\
&= \liminf_{\varepsilon\downarrow 0} R_{t^\varepsilon_1} R_{t^\varepsilon_2 - t^\varepsilon_1} \cdots R_{t - t^\varepsilon_{J(\varepsilon)-1}}(1_{\mathcal{Z}})(Z_0) = R_t(Z_0, \mathcal{Z})
\end{aligned}$$

donde la desigualdad se debe al Lema de Fatou, la segunda igualdad al Lema 1.2 y la última igualdad a que $(R_t)_{t\geq 0}$ es un semigrupo de transición. Considerando ahora al complemento de $\mathcal{Z}$ en $\mathcal{X}\times\mathcal{Y}$, se obtiene la desigualdad contraria, por lo que $P[Z_t \in \mathcal{Z} \mid Z_0] = R_t(Z_0, \mathcal{Z})$ y al ser $\mathcal{Z}$ arbitrario, se concluye que $(R_t)_{t\geq 0}$ es el semigrupo de transición de Z. □

El Teorema 2 de [6] es un cierto tipo de resultado opuesto al Teorema 1 del mismo. En este se identifica a los generadores de entrelazamientos de difusiones. Sean $\mathcal{A}^X$ y $\mathcal{A}^Y$ los generadores definidos previamente que siguen satisfaciendo la Condición 1 y X y Y las correspondientes difusiones. Asúmase que existe un proceso de Feller Z que satisface los puntos (i), (ii) y (v) en la Definición 3.2 y las siguientes dos hipótesis

(iii)' Para cualquier $t > 0$, dada Z^1_0, las variables aleatorias Z^2_0 y Z^1_t son condicionalmente independientes.

(iv)' Para cualquier $t > 0$, dada Z^2_t, las variables aleatorias Z^2_0 y Z^1_t son condicionalmente independientes.

Así, se tiene el siguiente

Teorema 3.14. *Supóngase que el generador de Z satisface la Condición 1 y el núcleo L satisface la Condición 2. Además, supóngase que*

(i) Λ_x está en el dominio de $\mathcal{A}^Y$ en $C_0(\mathcal{Y})$ para todo $x \in \mathcal{X}$,

(ii) la medida L_y pertenece al dominio de $(\mathcal{A}^X)^*$ para todo $y \in \mathcal{Y}$ y

(iii) para toda $x \in \mathcal{X}$, $\mathcal{A}^Y \Lambda_x$ es continua y existe una vecindad abierta acotada $U(x)$ de x tal que $\mathcal{A}^Y \Lambda_x$ está uniformemente acotada en $\mathcal{Y} \times U(x)$.

Entonces, el generador de Z está dado por (3.2) *y para cada función $f \in \mathcal{D}(\mathcal{A}^X) \cap C_0(\mathcal{X})$, la relación*

$$L\mathcal{A}^X f = \mathcal{A}^Y L f \tag{3.30}$$

se cumple. Más aún, para cada función $f \in \mathcal{A}^X$ con soporte compacto se cumple (3.3)*, es decir,*

$$\int_{\mathcal{X}} (\mathcal{A}^X f)(x)\Lambda(y,x)dx = \int_{\mathcal{X}} (\mathcal{A}^Y \Lambda_x)(y) f(x)dx, \quad y \in \mathcal{Y}$$

y esto se satisface para toda $f \in \mathcal{A}^X \cap C_0(\mathcal{X})$ en los $y \in \mathcal{Y}$ para los cuales $\mathcal{A}^Y \Lambda.(y) \in L^1(\mathcal{X})$ (respecto a la medida de Lebesgue).

Obsérvese que si $\mathcal{X}$ está acotado, entonces esta última condición se cumple para toda $y \in \mathcal{Y}$ debido a la condición (iii) del teorema.

Demostración. Sean f y g funciones medibles acotadas en $\mathcal{X}$ y $t \geq 0$. Por la condición (iii)', se tiene

$$\begin{aligned}
\mathbb{E}[f(Z_0^1)g(Z_t^1) \mid Z_0^2] &= \mathbb{E}[f(Z_0^1)\mathbb{E}[g(Z_t^1) \mid Z_0^1, Z_0^2] \mid Z_0^2] \\
&= \mathbb{E}[f(Z_0^1)\mathbb{E}[g(Z_t^1) \mid Z_0^1] \mid Z_0^2] \\
&= \mathbb{E}[f(Z_0^1)(P_t g)(Z_0^1) \mid Z_0^2] \\
&= L(f(P_t g))(Z_0^2)
\end{aligned}$$

y si h es una función medible y acotada en $\mathcal{Y}$, entonces, por la condición (iv)',

$$\begin{aligned}
\mathbb{E}[g(Z_t^1)h(Z_t^2) \mid Z_0^2] &= \mathbb{E}[h(Z_t^2)\mathbb{E}[g(Z_t^1) \mid Z_0^2, Z_t^2] \mid Z_0^2] \\
&= \mathbb{E}[h(Z_t^2)\mathbb{E}[g(Z_t^1) \mid Z_t^2] \mid Z_0^2] \\
&= \mathbb{E}[h(Z_t^2)(Lg)(Z_t^2) \mid Z_0^2] \\
&= Q_t(h(Lg))(Z_0^2)
\end{aligned}$$

Esto muestra que las ecuaciones (3.21) y (3.22) siguen cumpliéndose y se puede proceder como antes del Lema 3.12 para probar que el semigrupo de transición de Z tiene la forma dada por la ecuación (3.24). En particular, para cualquier función h en $C_0(\mathcal{X} \times \mathcal{Y})$,

$$\mathbb{E}[h(Z_t) \mid Z_0] = \int_{\mathcal{X}\times\mathcal{Y}} \frac{\Lambda(y_1, x_1)h(x_1, y_1)}{(Q_t\Lambda_{x_1})(Z_0^2)} P_t(Z_0^1, dx_1)Q_t(Z_0^2, dy_1) \tag{3.31}$$

y de nuevo, la distribución de Z_t es absolutamente continua respecto al producto de las leyes de $X(t)$ y $Y(t)$.

Por el Teorema de Stone-Weierstrass, el conjunto de funciones h en $\mathcal{X}\times\mathcal{Y}$ de la forma $h(x,y) = f(x)g(y)$ donde $f \in C_0^\infty(\mathcal{X})$ y $g \in C_0^\infty(\mathcal{Y})$ es denso en $C_0(\mathcal{X}\times\mathcal{Y})$. Por el Teorema (2.13), $\mathcal{A}$ es un operador cerrado y ya que $C_0^\infty(\mathcal{X}\times\mathcal{Y})$ es un centro para $\mathcal{A}$, solo hace falta identificar la acción de $\mathcal{A}$ en funciones $h(x,y) = f(x)g(y)$ donde f y g son como antes. Más aún, como el operador es local, podemos asumir que al fijar $x \in \mathcal{X}$, el soporte de f está contenido en $U(x)$.
Por hipótesis, para cualquier $x_1 \in \mathcal{X}$, Λ_{x_1} pertenece al dominio $\mathcal{A}^Y$. Por lo tanto, por el Lema 3.4 muestra que $\Lambda_{x_1} g$ pertenece al dominio de $\mathcal{A}^Y$. Haciendo $h(x,y) = f(x)g(y)$ en la ecuación (3.31), el Teorema 2.4 implica que

$$\begin{aligned}\mathbb{E}[f(Z_t^1)g(Z_t^2) \mid Z_0] = \\ \int_{\mathcal{X}} \frac{\Lambda(Z_0^2,x_1)g(Z_0^2) + t(\mathcal{A}^Y(g\Lambda_{x_1}))(Z_0^2) + t\varepsilon_1(t,x_1,Z_0^2)}{\Lambda(y_0,x_1) + t(\mathcal{A}^Y\Lambda_{x_1})(Z_0^2) + t\varepsilon_2(t,x_1,Z_0^2)} f(x_1)P_t(Z_0^1,dx_1)\end{aligned} \tag{3.32}$$

donde

$$\begin{aligned}\varepsilon_1(t,x_1,Z_0^2) &= \frac{1}{t}\int_0^t (Q_s\mathcal{A}^Y(g\Lambda_{x_1}))(Z_0^2) - (\mathcal{A}^Y(g\Lambda_{x_1}))ds \\ \varepsilon_2(t,x_1,Z_0^2) &= \frac{1}{t}\int_0^t (Q_s\mathcal{A}^Y\Lambda_{x_1})(Z_0^2) - (\mathcal{A}^Y\Lambda_{x_1})ds\end{aligned}$$

En vista de la ecuación (3.4) y la continuidad de Λ, $\nabla_y\Lambda$ y $\mathcal{A}^Y\Lambda$, las funciones $\mathcal{A}^Y(\Lambda g)$ y $\mathcal{A}^Y\Lambda$ están uniformemente acotadas en $\mathcal{Y}\times U(x)$ (recuérdese que $U(x)$ es acotado) por lo que ambos ε_1 y ε_2 convergen a cero cuando t tiende a 0 uniformemente en $x_1 \in U(x)$.
Aplicando la igualdad

$$\frac{a_1 + ta_2 + ta_3}{b_1 + tb_2 + tb_3} = \frac{a_1}{b_1} + t\frac{a_2b_1 - a_1b_2}{b_1^2} + t\frac{a_3b_1^2 - a_1b_1b_3 + t(a_1b_2^2 + a_1b_2b_3 - a_2b_1b_2 - a_2b_1b_3)}{b_1^3 + tb_1^2(b_2+b_3)} \tag{3.33}$$

a la fracción en (3.32), el lado derecho de ésta se puede escribir como

$$\int_{\mathcal{X}} \left(g(Z_0^2) + t\frac{(\mathcal{A}^Y(g\Lambda_{x_1}))(Z_0^2) - g(Z_0^2)(\mathcal{A}^Y\Lambda_{x_1})(Z_0^2)}{\Lambda(y_0,x_1)} + t\varepsilon_3(t,x_1,Z_0^2) \right) f(x_1)P_t(Z_0^1,dx_1) \tag{3.34}$$

donde una expresión explícita de ε_3 se puede obtener a partir de (3.33). Obsérvese que el numerador de ε_3 depende de Λ, Λg, $\mathcal{A}^Y\Lambda_x$, $\mathcal{A}^Y(g\Lambda_x)$, ε_1 y ε_2. Por continuidad de las primeras cuatro funciones y debido a que ε_1 y ε_2 convergen a 0 uniformemente en $x_1 \in U(x)$, el numerador converge también a cero uniformemente en x_1. Para t suficientemente pequeña, el denominador es estrictamente positivo y al ser $U(x)$ acotado, se sigue que el denominador está uniformemente acotado por abajo por un número positivo. Luego, ε_3 converge a 0 uniformemente en $x_1 \in U(x)$.
Expandiendo el producto en (3.32), se analizará el comportamiento de cada sumando cuando t tiende a 0:

1. Puesto que f pertenece a $\mathcal{D}(\mathcal{A}^X)$, se tiene

$$\int_{\mathcal{X}} g(Z_0^2) f(x_1) P_t(Z_0^1, dx_1) = f(Z_0^1) g(Z_0^2) + t(\mathcal{A}^X f)(Z_0^1) g(Z_0^2) + o(t)$$

2. Por la ecuación (3.4) y la continuidad en x_1 de las funciones involucradas,

$$\begin{aligned}
&\int_{\mathcal{X}} t \frac{(\mathcal{A}^Y(g\Lambda_{x_1}))(Z_0^2) - g(Z_0^2)(\mathcal{A}^Y \Lambda_{x_1})(Z_0^2)}{\Lambda(y_0, x_1)} f(x_1) P_t(Z_0^1, dx_1) \\
&= t \int_{\mathcal{X}} \frac{(\nabla_y \Lambda(Z_0^2, x_1))' \rho(Z_0^2) \nabla g(Z_0^2) + \Lambda(Z_0^2, x_1)(\mathcal{A}^Y g)(Z_0^2)}{\Lambda(Z_0^2, x_1)} f(x_1) P_t(Z_0^1, dx_1) \\
&= t \int_{\mathcal{X}} \left(\frac{(\nabla_y \Lambda(Z_0^2, x_1))'}{\Lambda(Z_0^2, x_1)} \rho(Z_0^2) \nabla g(Z_0^2) + (\mathcal{A}^Y g)(Z_0^2) \right) f(x_1) P_t(Z_0^1, dx_1) \\
&= t \Bigg((\nabla_y V(Z_0^2, Z_0^1))' \rho_(Z_0^2) \nabla g(Z_0^2) + (\mathcal{A}^Y g)(Z_0^2) \Bigg) f(Z_0^1) + o(t)
\end{aligned}$$

3. La convergencia uniforme en $x_1 \in U(x)$ de ε_3, muestra entonces que

$$\int_{\mathcal{X}} t\varepsilon_3(t, x_1, Z_0^2) f(x_1 = P_t(Z_0^1, dx_1) = o(t)$$

Todo esto es válido sin especificar la condición inicial Z_0. Al tomar como condición inicial $Z_0 = z_0 = (x_0, y_0)$, de lo anterior se obtiene que

$$\mathbb{E}_{z_0}[f(Z_t^1) g(Z_t^2)] = f(x_0) g(y_0) + \mathcal{A}^Z(fg)(z_0) + o(t)$$

donde $\mathcal{A}^Z$ está definido por (3.2). Esto implica entonces que el generador de Z en la función fg está dado por $\mathcal{A}^Z(fg)$ como se quería.
En este punto, solo hace falta probar que se cumple (3.3) y (3.30). Para obtener (3.30), sea $f \in C_0(\mathcal{X})$ una función en el dominio de $\mathcal{A}^X$. Usando la relación de entrelazamiento, se obtiene

$$\frac{LP_t - Lf}{t} = \frac{Q_t Lf - Lf}{t} \tag{3.35}$$

Como f está en el dominio de $\mathcal{A}^X$, se tiene $\frac{P_t f - f}{t} \to \mathcal{A}^X$ en $C_0(\mathcal{A}^X)$ y debido a la continuidad de L, se tiene entonces que el lado izquierdo de (3.35) converge a $L\mathcal{A}^X f$ en $C_0(\mathcal{Y})$. Así, el límite en $C_0(\mathcal{Y})$ del lado derecho de (3.35) existe, lo cual implica por definición que $Lf \in \mathcal{D}(\mathcal{A}^Y)$ y $\mathcal{A}^Y Lf = L\mathcal{A}^X f$ como se quería. Finalmente, para obtener (3.3), sea f una función en $\mathcal{A}^X$ con soporte compacto. Como los semigrupos $(P_t)_{t \geq 0}$ y $(Q_t)_{t \geq 0}$ están entrelazados, se tiene $LP_t f = Q_t Lf$ para cada $t \geq 0$. Por el Teorema de Fubini, esto es equivalente a

$$\int_{\mathcal{X}} \Lambda(y, x)(P_t f)(x) dx = \int_{\mathcal{X}} f(x)(Q_t \Lambda_x)(y) dx \quad y \in \mathcal{Y}, t \geq 0$$

de donde se obtiene

$$\int_{\mathcal{X}} \Lambda(y,x)\frac{(P_t f)(x) - f(x)}{t}dx = \int_{\mathcal{X}} f(x)\frac{(Q_t\Lambda_x)(y) - \Lambda_x(y)}{t}dx \quad y \in \mathcal{Y}, t \geq 0 \tag{3.36}$$

Ya que Λ_y es integrable respecto a la medida de Lebesgue en $\mathcal{X}$ y $\frac{P_t f - f}{t}$ converge uniformemente a $\mathcal{A}^X f$, por el Teorema de convergencia dominada, el lado izquierdo de 3.36 converge a $\int_{\mathcal{X}} (\mathcal{A}^X f)(x)\Lambda(y,x)dx$.
Asumiendo primero que el soporte de f está contenido en algún $U(x')$ para algún $x' \in \mathcal{X}$, el lado derecho de (3.36) se puede escribir como

$$\int_{\mathcal{X}} f(x)\frac{(Q_t\Lambda_x)(y) - \Lambda_x(y)}{t}dx = \int_{U(x')} f(x)(\mathcal{A}^Y\Lambda_x)(y) + f(x)\varepsilon_1(t,x,y)dx$$

donde ε_1 está definido como antes. Por las observaciones hechas anteriormente, ε_1 converge uniformemente a 0 en $U(x')$ y debido a que tanto f como $U(x')$ están acotados, una nueva aplicación del Teorema de convergencia dominada implica que $\frac{1}{t}\int_{\mathcal{X}} f(x)((Q_t\Lambda_x)(y) - \Lambda_x(y))dx \to \int_{\mathcal{X}} (\mathcal{A}^Y\Lambda_x)(y)f(x)dx$. Igualando ambos límites, se obtiene (3.3).
Si f tiene soporte compacto arbitrario, podemos cubrir el soporte de f con una cantidad finita de los $U(x)$, considerar las intersecciones con dichos $U(x)$ y proceder a hacer ajenos estos conjuntos, de manera que se obtiene una partición del soporte de f en los cuales se satisface (3.3). Debido a la linealidad de $\mathcal{A}^X$ y de la integral, dicha relación se sigue cumpliendo aún en este caso.
Si para algún $y \in \mathcal{Y}$ la función $(\mathcal{A}^Y\Lambda.)(y)$ es integrable (respecto a al medida de Lebesgue) y $f \in \mathcal{A}^X \cap C_0(\mathcal{X})$, se puede considerar una sucesión de funciones en $C_0^\infty(\mathcal{X})$, $(f_n)_{n\geq 0}$ tales que $f_n \to f$ y $\mathcal{A}^X f_n \to \mathcal{A}^X f$ en $C_0(\mathcal{X})$. Así,

$$\begin{aligned}\left|\int_{\mathcal{X}} \Lambda(y,x)(\mathcal{A}^X f)(x) - f(x)(\mathcal{A}^Y\Lambda_x)(y)dx\right| &\leq \left|\int_{\mathcal{X}} \Lambda(y,x)\Big((\mathcal{A}^X f)(x) - (\mathcal{A}^X f_n)(x)\Big)dx\right| \\ &\quad + \left|\int_{\mathcal{X}} (f_n(x) - f(x))(\mathcal{A}^Y\Lambda_x)(y)dx\right| \\ &\leq \|\mathcal{A}^X f - \mathcal{A}^X f_n\| \int_{\mathcal{X}} \Lambda(y,x)dx \\ &\quad + \|f_n - f\| \int_{\mathcal{X}} |(\mathcal{A}^Y\Lambda_x)(y)|dx\end{aligned}$$

y debido a la integrabilidad de Λ_y y $(\mathcal{A}^Y\Lambda.)(y)$, la última expresión converge a 0, lo cual concluye con la prueba. $\square$

CAPÍTULO 4

PROPIEDADES Y EJEMPLOS

Este capítulo está dividido en dos secciones. En la primer sección se demostrarán algunos teoremas relacionados con difusiones, así como algunas de sus propiedades y generalizaciones del teorema 3.3. En la segunda sección se enunciarán ejemplos de difusiones entrelazadas, así como procesos en los que los semigrupos están entrelazados según la Definición 3.2, en los cuales se discutirá la posible aplicación de los teoremas 3.3 y 3.14.

Propiedades de los entrelazamientos

En el teorema 3.3 se impuso la condición de que Λ_y fuera una medida de probabilidad para cada $y \in \mathcal{Y}$ y que cumpliera la ecuación 3.3. Supóngase que solo se pide que la ecuación (3.3) sea satisfecha y que Λ_y sea densidad de una medida positiva y finita con masa total $\tau(y) = \int_{\mathcal{X}} \Lambda(y,x)dx$. Se puede definir una densidad normalizada

$$\xi(y,x) = \frac{\Lambda(y,x)}{\tau(y)}$$

y un operador de transición Ξ correspondiente a ξ. El teorema 4.1 muestra que Ξ es enlace de un entrelazamiento entre el semigrupo $(P_t)_{t\geq 0}$ y un semigrupo obtenido a partir de $(Q_t)_{t\geq 0}$.

Teorema 4.1. *Supóngase que el proceso Y satisface la condición 1 y tiene condición inicial $Y_0 = y_0$. Sea $(\mathscr{F}_t^Y)_{t\geq 0}$ la filtración natural aumentada generada por Y. Asúmase que la variación total de la medida $(\mathcal{A}^X)^* L_y$ está localmente acotada al variar y y que la función τ es continua. Entonces, el proceso $(\tau(Y_t))_{t\geq 0}$ es una semi martingala local positiva respecto a $(\mathscr{F}_t^Y)_{t\geq 0}$.*
Sea $\tilde{Y}$ un proceso de Feller cuyo generador $\mathcal{A}^\tau$ satisface

$$\mathcal{A}^\tau \phi = \tau^{-1}\mathcal{A}^Y(\tau\phi) \tag{4.1}$$

para funciones ϕ para las cuales $\tau\phi \in \mathcal{D}(\mathcal{A}^Y)$. Entonces, su semigrupo $(Q_t^\tau)_{t\geq 0}$ satisface $Q^\tau\langle\Xi\rangle P$.

Por la ecuación 3.3, la variación total de la medida $(\mathcal{A}^X)^* L_y$ está localmente acotada al variar y si y solo si la función $\int_{\mathcal{X}} |\mathcal{A}^Y \Lambda_x| dx$ está localmente acotada.
Obsérvese que la ecuación (4.1) indica que el proceso $\tilde{Y}$ es en realidad la transformada h de Doob de Y.

Demostración. Para $R > 0$, defínase

$$v_R = \text{ínf}\{t \geq 0 \mid Y_t \notin B_R(y_0)\} \tag{4.2}$$

donde $B_r(y_0)$ es la bola abierta de radio R alrededor de y_0. Por continuidad de Y, cada v_R es un tiempo de paro y $v_R \to \infty$ cuando $R \to \infty$, por lo que para demostrar el lema, solo hace falta ver que el proceso $(\tau(Y_{t\wedge v_R})_{t\geq 0}$ es una martingala. Sean $0 \leq s \leq t$ y $A \in \mathscr{F}_s$. Hay que probar

$$\mathbb{E}[\tau(Y_{t\wedge v_R})1_A] = \mathbb{E}[\tau(Y_{s\wedge v_R})1_A] \tag{4.3}$$

Sea $(f_n)_{n\in\mathbb{N}}$ una sucesión en $C_0(\mathcal{X})$ de funciones crecientes no negativas que convergen a la función constante 1 y para cada n, defínase $g_n = \int_0^1 P_s f_n ds$. Ahora,

$$\begin{aligned}
\frac{P_t g_n - g_n}{t} &= \frac{1}{t}\int_t^{t+1} P_s f_n ds - \frac{1}{t}\int_0^1 P_s f_n ds \\
&= \frac{1}{t}\int_1^{t+1} P_s f_n ds - \frac{1}{t}\int_0^t P_s f_n ds \\
&= \frac{1}{t}\int_0^{t} P_{1+s} f_n ds - \frac{1}{t}\int_0^t P_s f_n ds \\
&= P_1\Big(\frac{1}{t}\int_0^{t} P_s f_n ds\Big) - \frac{1}{t}\int_0^t P_s f_n ds
\end{aligned}$$

Esta última expresión converge a $P_1 f_n - f_n$ cuando $t \to 0$, lo cual muestra que g_n está en el dominio de $\mathcal{A}^X$ y $\mathcal{A}^X g_n = P_1 f_n - f_n$. Por el teorema de Fubini y el teorema de la convergencia dominada, se tiene

$$\begin{aligned}
\mathbb{E}[\tau(Y_{t\wedge v_R})1_A] - \mathbb{E}[\tau(Y_{s\wedge v_R})1_A] &= \int_{\mathcal{X}} \mathbb{E}[\Lambda(Y_{t\wedge v_R}, x)1_A - \Lambda(Y_{s\wedge v_R}, x)1_A]dx \\
&= \lim_{n\to\infty} \int_{\mathcal{X}} \mathbb{E}[\Lambda(Y_{t\wedge v_R}, x)1_A - \Lambda(Y_{s\wedge v_R}, x)1_A] g_n(x) dx
\end{aligned} \tag{4.4}$$

La fórmula de Dynkin implica que el término dentro de la esperanza se puede escribir como

$$\Lambda(Y_{t\wedge v_R}, x)1_A - \Lambda(Y_{s\wedge v_R}, x)1_A = 1_A \int_{s\wedge v_R}^{t\wedge v_R} \mathcal{A}^Y \Lambda_x(Y_u) du$$

Para cada n, se sigue que $|g_n(\mathcal{A}^Y\Lambda_x)(Y_{u\wedge v_R})| \leq \|g_n\| |(\mathcal{A}^Y\Lambda_x)(Y_{u\wedge v_R})|$ y por hipótesis, esta función es integrable en x. Luego, (4.4) se puede escribir como

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}\left[1_A \int_{s\wedge v_R}^{t\wedge v_R} \int_{\mathcal{X}} \mathcal{A}^Y \Lambda_x(Y_u) g_n(x) dx du\right] &= \lim_{n\to\infty} \mathbb{E}\left[1_A \int_{s\wedge v_R}^{t\wedge v_R} \int_{\mathcal{X}} \Lambda_x(Y_u) \mathcal{A}^X g_n(x) dx du\right] \\
&= \lim_{n\to\infty} \mathbb{E}\left[1_A \int_{s\wedge v_R}^{t\wedge v_R} \int_{\mathcal{X}} \Lambda_x(Y_u)(P_1 f_n - f_n)(x) dx du\right]
\end{aligned}$$

Por el teorema de convergencia monótona, $Pf_n - f_n \to 0$ puntualmente en $\mathcal{X}$ y ya que P_1 es una contracción, entonces $|P_1 f_n(x) - f_n(x)| \leq 2$, de donde

$$|\Lambda_x(Y_u)(P_1 f_n - f_n)(x)| \leq 2\Lambda(Y_u, x). \tag{4.5}$$

Esta última función es integrable en $\mathcal{X}$, por lo que el teorema de convergencia dominada implica que

$$\int_{\mathcal{X}} \Lambda_x(Y_u)(P_1 f_n - f_n)(x)dx \to 0$$

para cada $u \in [s \wedge v_R, t \wedge v_R]$. Por la ecuación (4.5), se obtiene

$$1_{[s\wedge v_r, t\wedge v_r]}(u)\left|\int_{\mathcal{X}} \Lambda_x(Y_u)(P_1 f_n - f_n)(x)dx\right| \leq 2\tau(Y_u)1_{[s\wedge v_r, t\wedge v_r]}(u)$$

y como τ es continua, la expresión de la derecha está acotada para cualquier trayectoria en el intervalo $[0, t]$, por lo que a partir de una aplicación del teorema de convergencia domina, se concluye que

$$\mathbb{E}[\tau(Y_{t\wedge v_R})1_A] - \mathbb{E}[\tau(Y_{s\wedge v_R})1_A] = \lim_{n\to\infty} \mathbb{E}\left[1_A \int_{s\wedge v_R}^{t\wedge v_R} \int_{\mathcal{X}} \Lambda_x(Y_u)(P_1 f_n - f_n)(x)dxdu\right] = 0.$$

Para demostrar la última afirmación, obsérvese que para cualquier función $f\mathcal{D}(\mathcal{A}^X)$ y $y \in \mathcal{Y}$, la ecuación 3.3 implica

$$\begin{aligned}\int_{\mathcal{X}} f(x)(\mathcal{A}^\tau \xi_x)(y)dx &= \int_{\mathcal{X}} \tau(y)^{-1} f(x)(\mathcal{A}^Y \Lambda_x)(y)dx \\ &= \int_{\mathcal{X}} \tau(y)^{-1}\Lambda(y,x)(\mathcal{A}^X f)(x)dx = \int_{\mathcal{X}} \xi(y,x)(\mathcal{A}^X f)(x)dx\end{aligned}$$

por lo que el enunciado se sigue por un procedimiento similar al de la demostración del Lema 3.8 (con Ξ, ξ, $(Q_t^\tau)_{t\geq 0}$, X y $\tilde{Y}$ en lugar de L, Λ, $(Q_t)_{t\geq 0}$, Z^1 y Z^2 respectivamente). □

Supóngase que $\tilde{Y}$ es un proceso tal que para cada $t \geq 0$, la distribución de $\tilde{Y}$ es absolutamente continua respecto a la distribución de Y en $\mathscr{F}_t^Y$ y la densidad correspondiente está dada por $\tau(y_0)^{-1}\tau(Y_t)$ (esto es posible si, por ejemplo, la martingala local $(\tau(Y_t))_{t\geq 0}$ es en realidad una martingala). Así, se tiene el siguiente

Proposición 4.2. *Supóngase que τ está acotada inferiormente por $\alpha > 0$. Entonces, el proceso $\tilde{Y}$ es un proceso de Feller cuyo generador es $\mathcal{A}^\tau$.*

Demostración. Para demostrar la propiedad de Markov es suficiente demostrar que para $h \in C_0^\infty$ y $0 \leq s < t$,

$$\mathbb{E}[h(\tilde{Y}_t) \mid \mathscr{F}_s^Y] = \tau(\tilde{Y}_s)^{-1} Q_{t-s}(\tau h)(\tilde{Y}_s) \tag{4.6}$$

Con este fin, sea $A \in \mathscr{F}_s^Y$. Por hipótesis,

$$\begin{aligned}\mathbb{E}[\tau(\tilde{Y}_s)^{-1} Q_{t-s}(\tau h)(\tilde{Y}_s) 1_A] &= \frac{1}{\tau(y_0)} \mathbb{E}[Q_{t-s}(\tau h)(Y(s)) 1_A] \\ &= \mathbb{E}[\mathbb{E}[(\tau h)(Y_t) 1_A \mid \mathscr{F}_s^Y]] = \mathbb{E}[h(\tilde{Y}_t) 1_A].\end{aligned}$$

Para demostrar la propiedad de Feller, considérese la aplicación $y \mapsto \tau(y)^{-1} Q_t(\tau h)(y)$ para alguna $h \in C_0(\mathcal{Y})$ y $0 \leq t$, cuya pertenencia a $C_0(\mathcal{Y})$ hay que demostrar. Aproximando uniformemente a h por funciones infinitamente diferenciables con soporte compacto, se puede asumir que $h \in C_0^\infty(\mathcal{Y})$. Así, $\tau h \in C_0(\mathcal{Y})$ por lo que la continuidad de $y \mapsto \tau(y)^{-1} Q_t(\tau h)(y)$ se debe a la propiedad de Feller para Y. Luego,

$$\begin{aligned}|\tau(y_0)^{-1} Q_t(\tau h)(y_0)| &= \left| \mathbb{E}\big[\tau(y_0)^{-1} \mathbb{E}[\tau(Y_t) h(Y_t) \mid \mathscr{F}_{v_R}^Y]\big] \right| \\ &\leq \frac{\sup\{\tau(x) | x \in \operatorname{supp}(h)\}}{\alpha^{-1}} \mathbb{E}[|h(Y_t|] = \frac{\sup\{\tau(x) | x \in \operatorname{supp}(h)\}}{\alpha^{-1}} Q_t(|h|)(y_0)\end{aligned}$$

Y debido a que $(Q_t)_{t \geq 0}$ es de Feller, la expresión del final converge a cero cuando $y_0 \to \infty$ y por la Definición 2.2, el semigrupo de $\tilde{Y}$ es de Feller. Finalmente, la fórmula para el generador se sigue de la ecuación (4.6). □

Para obtener otra extensión del teorema 3.3 se puede considerar una extensión del diagrama presentado al inicio del capítulo anterior. Esto es, supóngase que se tienen tres procesos de Markov Z^1, Z^2 y Z^3 con semigrupos de transición $(P_t)_{t\geq 0}$, $(Q_t)_{t\geq 0}$ y $(R_t)_{t\geq 0}$ respectivamente de manera que los semigrupos satisfacen $R\langle \tilde{L} \rangle Q$ y $Q\langle L \rangle P$ para ciertos operadores estocásticos de transición entre los espacios de estados correspondientes. Consideraciones parecidas a las del inicio del capítulo anterior sugieren el siguiente diagrama

$$\begin{tikzcd}
Z_t^3 \arrow[r, "R_s"] \arrow[d, "\tilde{L}"'] & Z_{t+s}^3 \arrow[d, "\tilde{L}"] \\
Z_t^2 \arrow[r, "Q_s"] \arrow[d, "L"'] & Z_{t+s}^2 \arrow[d, "L"] \\
Z_t^1 \arrow[r, "P_s"] & Z_{t+s}^1
\end{tikzcd}$$

El siguiente teorema muestra que este diagrama se puede obtener al considerar una iteración del acoplamiento construido en el teorema 3.3.

Teorema 4.3. *Supóngase las hipótesis del teorema 3.3. Sea S otra difusión con espacio de estados $\mathcal{S} \subset \mathbb{R}^k$ y generador*

$$\mathcal{A}^S = \sum_{i=1}^{k} \eta_i \frac{\partial}{\partial s_i} + \frac{1}{2} \sum_{i,j=1}^{k} \sigma_{ij} \frac{\partial^2}{\partial s_i \partial s_j} \tag{4.7}$$

el cual satisface la Condición 1. Asúmase que $\tilde{L}$ manda las funciones en $C_0(\mathcal{Y})$ a funciones en $C_0(\mathcal{S})$ y es el operador inducido por el $\tilde{\Lambda}$, el cual es estrictamente positivo y defínase $\tilde{V} = \log \tilde{\Lambda}$. Considérese la difusión $Z = (Z^1, Z^2, Z^3)$ con estado de espacios $\mathcal{X} \times \mathcal{Y} \times \mathcal{S}$ y generador

$$\mathcal{A}^Z = \mathcal{A}^X + \mathcal{A}^Y + \mathcal{A}^S + (\nabla_y V)' \rho \nabla_y + (\nabla_s \tilde{V})' \sigma \nabla_s$$

Supóngase que la densidad condicional de Z_0^2 dado Z_0^3 es $\tilde{\Lambda}$ y que la densidad condicional de Z_0^1 dados Z_0^2 y Z_0^3 es Λ (por lo que, en particular, condicionado a Z_0^2, Z_0^1 es independiente de Z_0^3). Si $\tilde{\Lambda}$ es tal que

(i) Para cada $y \in \mathcal{Y}$, $\tilde{\Lambda}_y$ está en el dominio de $\mathcal{A}^S$ y $\mathcal{A}^S \tilde{\Lambda}$ es continua en $\mathcal{S} \times \mathcal{Y}$ y acotada en $\mathcal{S} \times K$ para cualquier subconjunto compacto $K \subset \mathcal{Y}$,

(ii) Para $s \in \mathcal{S}$, $\tilde{L}_s$ está en el dominio de $(\mathcal{A}^Y)^*$ y

(iii) la medida $(\mathcal{A}^Y)^* \tilde{L}_s$ es absolutamente continua y su densidad está dada por $(\mathcal{A}^S \tilde{\Lambda}_y)$.

entonces, $Z = S\langle L \rangle (Z^1, Z^2)$

Demostración. Por la forma del generador de Z, al aplicar la fórmula de Itô a funciones de (Z^1, Z^2) se ve que (Z^1, Z^2) resuelve el problema de la martingala asociado con el generador 3.2, por lo que el teorema 3.3 implica que que (Z^1, Z^2) es el entrelazamiento construido en dicho teorema. Sea $\mathcal{A}^{Z^1,Z^2}$ el generador correspondiente.
Usando las hipotesis acerca de las distribuciones condicionales de Z_0^1, Z_0^2 y Z_0^3 y que los generadores $\mathcal{A}^{Z^1,Z^2}$ y $\mathcal{A}^S$ satisfacen la Condición 1, se obtiene que las difusiones (Z^1, Z^2) y S satisfacen nuevamente las hipótesis del teorema 3.3 y si L' es el operador definido por

$$L'f(s) = \int_{\mathcal{X}\times\mathcal{Y}} \tilde{\Lambda}(s,y)\Lambda(y,x)f(x,y)dxdy$$

para f una función medible y acotada en $\mathcal{X}\times\mathcal{Y}$, entonces, por el teorema de Fubini, L satisface la Condición 2, por lo que para poder aplicar el teorema 3.3 solo hace falta verificar que la medida L'_s se encuentra en el dominio de $(\mathcal{A}^{Z^1,Z^2})^*$ y la medida $(\mathcal{A}^{Z^1,Z^2})^*L_s$ tiene densidad dada por $\mathcal{A}^S(\tilde{\Lambda}_y)\Lambda$. Seún la observación hecha al final del enunciado del teorema 3.3, esto es equivalente a

$$\int_{\mathcal{X}\times\mathcal{Y}} (\mathcal{A}^{Z^1,Z^2}f)(x,y)\tilde{\Lambda}(s,y)\Lambda(y,x)dxdy = \int_{\mathcal{X}\times\mathcal{Y}} (\mathcal{A}^S\tilde{\Lambda}_y)(s)\Lambda(y,x)f(x,y)dxdy$$

para toda $f \in \mathcal{D}(\mathcal{A}^{Z^1,Z^2})$ y $s \in \mathcal{S}$. Sin embargo, por el punto (iii) del teorema, $\mathcal{A}^S\tilde{\Lambda}_y$ es la densidad de la medida $(\mathcal{A}^Y)^*\tilde{L}_s$, por lo que

$$\begin{aligned}\int_{\mathcal{X}\times\mathcal{Y}} (\mathcal{A}^{Z^1,Z^2}f)(x,y)\tilde{\Lambda}(s,y)\Lambda(y,x)dxdy &= \int_{\mathcal{X}\times\mathcal{Y}} f(x,y)\Lambda(y,x)d(\mathcal{A}^Y)^*\tilde{L}_s \\ &= \int_{\mathcal{X}\times\mathcal{Y}} \mathcal{A}^Y(f_x\Lambda_x)(y)\tilde{\Lambda}(s,y)dxdy\end{aligned} \tag{4.8}$$

y esta es la ecuación que se demostrará. Usando que el operador $\mathcal{A}^{Z^1,Z^2}$, un argumento de aproximación muestra que podemos asumir que $f \in C_0^\infty(\mathcal{X}\times\mathcal{Y})$. Así, por el Lema 3.4, se tiene

$$\begin{aligned}\int_{\mathcal{X}\times\mathcal{Y}} \mathcal{A}^Y(f_x\Lambda_x)(y)\tilde{\Lambda}(s,y)dxdy &= \int_{\mathcal{X}\times\mathcal{Y}} \tilde{\Lambda}(s,y)\Big[\Lambda_x(\mathcal{A}^Y f_x) + (\nabla_y\Lambda)'\rho(\nabla_y f) + f_x(\mathcal{A}^Y\Lambda_x)\Big](x,y)dxdy \\ &= \int_{\mathcal{X}\times\mathcal{Y}} \tilde{\Lambda}(s,y)f(x,y)(\mathcal{A}^Y\Lambda_x)(x,y)dxdy \\ &\quad + \int_{\mathcal{X}\times\mathcal{Y}} \tilde{\Lambda}(s,y)\big[\Lambda_x(\mathcal{A}^Y f_x) + (\nabla_y\Lambda)'\rho(\nabla_y f)\big](x,y)dxdy\end{aligned} \tag{4.9}$$

Por el teorema de Fubini y la ecuación 3.3, la primer integral en la última integral es igual a

$$\int_{\mathcal{Y}} \tilde{\Lambda}(s,y)\int_{\mathcal{X}} \Lambda(y,x)(\mathcal{A}^X f_y)(x)dxdy = \int_{\mathcal{X}\times\mathcal{Y}} \tilde{\Lambda}(s,y)\Lambda(y,x)(\mathcal{A}^X f_y)(x)dxdy$$

y al sustituir esta expresión en la última igualdad de (4.9) se obtiene (4.8). □

Observación. Es claro que, cuando las hipótesis sean satisfechas, este teorema se puede utilizar para crear acoplamientos $(Z^1, Z^2, \ldots, Z^d)$ usando cualquier número de difusiones.

El siguiente resultado podría ser interpretado como una regla de Bayes para entrelazamientos. Supóngase que $Q\langle L\rangle P$ para algunos semigrupos $(P_t)_{t\geq 0}$, $(Q_t)_{t\geq 0}$ y un núcleo de transición L. Se quiere saber bajo qué condiciones existe otro núcleo de transición $\widehat{L}$ de manera que $P\langle\widehat{L}\rangle Q$. Antes de enunciar un teorema que responda esto, se dará una definición preliminar

Definición 4.4. Se dice que dos semigrupos $(P_t)_{t\geq 0}$ y $(\hat{P}_t)_{t\geq 0}$ en $\mathbb{R}^d$ está en dualidad con respecto a una medida de probabilidad ν si se satisface

$$\int_{\mathbb{R}}^{d} (P_t f) g d\nu = \int_{\mathbb{R}}^{d} f(\hat{P}_t g) d\nu \quad \text{para toda } f \text{ y } g \text{ medibles y acotadas y todo } t \geq 0 \tag{4.10}$$

Se dice que $(P_t)_{t\geq 0}$ es reversible con respecto a ν si lo anterior se cumple con $(\hat{P}_t)_{t\geq 0} = (P_t)_{t\geq 0}$.

La definición puede enunciarse de la siguiente manera: el proceso de Markov con semigrupo $(P_t)_{t\geq 0}$ y distribuciónn inicial ν, es de nuevo un proceso de Markov al mirarse hacia atrás en el tiempo y su semigrupo de transición es $(\hat{P}_t)_{t\geq 0}$.
Considérense los semigrupos $(P_t)_{t\geq 0}$ y $(Q_t)_{t\geq 0}$ de dos difusiones como en la Condición 1 y un operador de transición estocástico L tal que $Q\langle L\rangle P$. Supóngase que existen semigrupos $(\hat{P}_t)_{t\geq 0}$ y $(\hat{Q}_t)_{t\geq 0}$ y dos medidas de probabilidad ν_1 y ν_2 tales que

(i) Los semigrupos $(P_t)_{t\geq 0}$ y $(\hat{P}_t)_{t\geq 0}$ están en dualidad con respecto a ν_1 y $(Q_t)_{t\geq 0}$ y $(\hat{Q}_t)_{t\geq 0}$ están en dualidad con respecto a ν_2.

(ii) El soporte de las medidas ν_1 y ν_2 son los conjuntos $\mathcal{X}$ y $\mathcal{Y}$ y son absolutamente continuas con respecto a la medida de Lebesgue con funciones de densidad h_1 y h_2 respectivamente.

(iii) Para cualesquiera medidas de probabilidad μ y ν en $\mathcal{X}$ y $\mathcal{Y}$ respectivamente, se tiene las convergencias débiles

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \mu P_s ds = \nu_1 \text{ y } \lim_{t\to\infty} \frac{1}{t}\int_0^t \nu Q_s ds = \nu_2 \tag{4.11}$$

Teorema 4.5. *Sea Λ la densidad correspondiente a L y supóngase L manda funciones continuas y acotadas en $\mathcal{X}$ a funciones continuas y acotadas en $\mathcal{Y}$. Defínase*

$$\hat{\Lambda}(x,y) = \Lambda(y,x)\frac{h_2(y)}{h_1(x)} \tag{4.12}$$

y sea $\hat{L}$ el operador de transición correspondiente. Entonces,

(i) El operador $\hat{L}$ es estocástico y $\hat{P}\langle\hat{L}\rangle\hat{Q}$.

(ii) Supóngase que las hipótesis del teorema 3.3 son satisfechas para cada una de las tercias $(P_t)_{t\geq 0}$, $(Q_t)_{t\geq 0}$, L *y* $(\hat{P}_t)_{t\geq 0}$, $(\hat{Q}_t)_{t\geq 0}$, $\hat{L}$ *y* $(R_t)_{t\geq 0}$ *y* $(\hat{R}_t)_{t\geq 0}$ *son los semigrupos correspondientes a los entrelazamientos* $Q\langle L\rangle P$ *y* $\hat{P}\langle\hat{L}\rangle\hat{Q}$ *construidos a partir del teorema 3.3 respectivamente.*
Sea ν_3 *la medidad de probabilidad en* $\mathcal{X}\times\mathcal{Y}$ *con densidad dada por* Λh_2*. Entonces, los semigrupos* $(R_y)_{t\geq 0}$ *y* $(\hat{R}_t)_{t\geq 0}$ *están en dualidad con respecto a* ν_3*.*

Observemos que si Λ es uniformemente continua o continua y acotada, entonces, por el teorema de convergencia dominada, L automáticamente satisface las hipótesis del teorema.

Demostración. Primero se probará que el operador $\hat{L}$ es estocástico, lo cual, por su definición, es equivalente a mostrar que

$$\int_{\mathcal{Y}}\Lambda(y,x)h_2(y)dy = h_1(x) \quad \text{para cada } x\in\mathcal{X} \tag{4.13}$$

Para ver esto, sea f una función continua y acotada en $\mathcal{X}$, ν una medida de probabilidad en $\mathcal{Y}$ y $\mu=\nu L$. Así, usando que $Q\langle L\rangle P$ y el teorema de Fubini,

$$\begin{aligned}\int_{\mathcal{X}} f(x)h_1(x)dx = \nu_1(f) &= \lim_{t\to\infty}\frac{1}{t}(\mu P_t)(f)\\ &= \lim_{t\to\infty}\frac{1}{t}(\nu Q_t)(Lf)\\ &= \nu_2(Lf)\\ &= \int_{\mathcal{Y}}(Lf)(y)h_2(y)dy = \int_{\mathcal{X}} f(x)\int_{\mathcal{Y}}\Lambda(y,x)h_2(y)dydx\end{aligned}$$

Usando la densidad de las funciones continuas y acotadas, se sigue la ecuación (4.13).
Para probar que $\hat{P}\langle\hat{L}\rangle\hat{Q}$, considérese funciones continuas y acotadas f y g en $\mathcal{X}$ y $\mathcal{Y}$ respectivamente. Para $t>0$ fijo, la relación (4.10), el teorema de Fubini y $Q\langle L\rangle P$ implican

$$\begin{aligned}\int_{\mathcal{X}}(\hat{P}_t\hat{L}g)f(x)\nu_1(dx) &= \int_{\mathcal{X}}(\hat{L}g)(x)(P_tf)(x)h_1(x)dx\\ &= \int_{\mathcal{X}}\Big(\int_{\mathcal{Y}}\Lambda(y,x)g(y)h_2(y)dy\Big)(P_tf)(x)dx\\ &= \int_{\mathcal{Y}}\Big(\int_{\mathcal{X}}\Lambda(y,x)(P_tf)(x)dx\Big)g(y)h_2(y)dy \qquad (4.14)\\ &= \int_{\mathcal{Y}}(LP_tf)(y)g(y)h_2(y)dy\\ &= \int_{\mathcal{Y}}(Q_tLf)(y)g(y)\nu_2(dy)\end{aligned}$$

Por otro lado, cálculos similares muestran que

$$\begin{aligned}\int_{\mathcal{X}}(\hat{L}\hat{Q}_t g)(x)f(x)\nu_1(dx) &= \int_{\mathcal{X}}\Big(\int_{\mathcal{Y}}\Lambda(y,x)(\hat{Q}_t g)(y)\nu_2(dy)\Big)f(x)dx\\ &= \int_{\mathcal{X}}\Big(\int_{\mathcal{Y}}(Q_t\Lambda_x)(y)g()\nu_2(dy)\Big)f(x)dx\\ &= \int_{\mathcal{Y}}\Big(\int_{\mathcal{X}}(Q_t\Lambda_x)(y)f(x)dx\Big)g(y)\nu_2(dy)\\ &= \int_{\mathcal{Y}}(Q_t Lf)(y)g(y)\nu_2(dy)\end{aligned} \tag{4.15}$$

Por lo tanto, la primer expresión en (4.14) es igual a la primer expresión en (4.15). Al ser f arbitraria, se sigue que $(\hat{P}_t\hat{L}g)h_2 = (\hat{L}\hat{Q}_t g)h_2$ casi seguramente respecto a la medida de Lebesgue. Luego, debido que h_2 es casi seguramente positiva y a la continuidad, se concluye que $\hat{P}_t\hat{L}g = \hat{L}\hat{Q}_t g$, por lo que $\hat{P}_t\hat{L}g = \hat{L}\hat{Q}_t$. Finalmente, solo hace falta mostrar que los semigrupos $(R_t)_{t\geq 0}$ y $(\hat{R}_t)_{t\geq 0}$ están en dualidad con respecto a ν_3. La ecuación (3.24) da una expresión para el primer semigrupo y, de manera análoga, se obtiene una expresión análoga para el semigrupo $(\hat{R}_t)_{t\geq 0}$ en términos de los semigrupos $(\hat{P}_t)_{t\geq 0}$ y $(\hat{Q}_t)_{t\geq 0}$.
Considérese el entrelazamiento $Z^2\langle L\rangle Z^1$ con distribución inicial ν_3 y sea $t > 0$ fijo. Puesto que Z_0^2 tiene distribución ν_2 se sigue, por la condición (i) de la Definición 3.2 que Z_t^2 tiene también distribución ν_2. Puesto que condicionado a Z_t^2, Z_t^1 tiene distribución dada por L se tiene entonces

$$\begin{aligned}P[Z_t^1\in A, Z_t^2\in B] &= \mathbb{E}[1_B(Z_t^2)L(Z_t^2,A)]\\ &= \int_B L(y,A)h_2(y)dy = \int_A\int_B\Lambda(y,x)h_2(y)dydx\end{aligned} \tag{4.16}$$

donde A y B son borelianos de $\mathcal{X}$ y $\mathcal{Y}$ respectivamente. Luego, la distribución conjunta de (Z_t^1, Z_t^2) está dada también por ν_3. Por lo tanto, ν_3 es una distribución invariante de (Z^1, Z^2). Si se demuestra que dado (Z_t^1, Z_t^2), la distribución condicional de (Z_0^1, Z_t^2) está dada por $(\hat{R}_t)_{t\geq 0}$, el Lema 1.2 implica

$$\begin{aligned}\int_{\mathbb{R}}^d (R_t f)g d\nu_3 &= \mathbb{E}[g(Z_0^1, Z_0^2)f(Z_t^1, Z_t^2)]\\ &= \mathbb{E}[f(Z_t^1, Z_t^2)\mathbb{E}[g(Z_0^1, Z_0^2)\mid Z_t^1, Z_t^2]]\\ &= \mathbb{E}[f(Z_t^1, Z_t^2)(\hat{R}_t g)(Z_t^1, Z_t^2)]\\ &= \int_{\mathbb{R}}^d f(\hat{R}_t g)d\nu_3\end{aligned}$$

por lo que la dualidad deseada se seguirá al demostrar que $(\hat{R}_t)_{t\geq 0}$ da dicha distribución condicional. Dénotese por λ a la distribución condicional de (Z_0^1, Z_0^2) dado (Z_t^1, Z_t^2). Se quiere demostrar que

$$\Lambda(x,y,A\times B) = \int_A\int_B \frac{\hat{\Lambda}(x',y')}{(\hat{P}_t\hat{\Lambda}_{y'})(x)}\hat{Q}_t(y,dy')\hat{P}_t(x,dx') \tag{4.17}$$

Para demostrar (4.17) se procederá de manera análoga a cuando se obtuvo una expresión para $(R_t)_{t\geq 0}$ antes del Lema 3.12.
Para conjuntos borelianos A, B, C y D en la σ-álgebra adecuada,

$$P[Z_0^1 \in A, Z_0^2 \in B, Z_t^1 \in C, Z_t^2 \in D] = \int_{C\times D} \lambda(x, y, A\times B)\nu_3(dx \times dy)$$

de donde se sigue que $\lambda(\cdot, A \times B)$ es la derivada de Radon Nikodym de la medida $P[Z_0^1 \in A, Z_0^2 \in B, (Z_t^1, Z_t^2) \in \cdot]$ respecto a ν_3.
Para cualquier función medible y acotada f en $\mathcal{X}$ y $y' \in \mathcal{Y}$, se tiene

$$\begin{aligned}
\int_{\mathcal{X}} f(x)(Q_t\Lambda_x)(y') &= (Q_t L)(f)(y') \\
&= (LP_t)(f)(y') \\
&= \int_{\mathcal{X}} (P_t f)(x)\Lambda(y', x)dx \\
&= \frac{1}{h_2(y')}\int_{\mathcal{X}} (P_t f)(x)\hat{\Lambda}(x, y')\nu_1(dx) \\
&= \frac{1}{h_2(y')}\int_{\mathcal{X}} f(x)(\hat{P}_t\hat{\Lambda}_{y'})(x)h_1(x)dx
\end{aligned}$$

Al ser f arbitraria, se concluye que $(Q_t\Lambda_.)(y') = \frac{h_1}{h_2(y')}\hat{P}_t\hat{\Lambda}_{y'}$ casi seguramente respecto a la medida de Lebesgue en $\mathbb{R}^m$. Luego, por el Lema 1.2, con $E = A\times B$ y $F = C \times D$,

$$\begin{aligned}
P[Z(0) \in E, Z(t) \in F] &= \int_A\int_B h_2(y)\Lambda(y, x)R_t(x, y, C\times D)dydx \\
&= \int_A\int_B\int_C\int_D \frac{h_2(y)\Lambda(y, x)\Lambda(y', x')}{(Q_t\Lambda_{x'})(y)}Q_t(y, dy')P_t(x, dx')dydx \\
&= \int_A\int_B\int_C\int_D \frac{h_2(y)^2\Lambda(y, x)\Lambda(y', x')}{h_1(x')(\hat{P}_t\hat{\Lambda}_y)(x')}Q_t(y, dy')P_t(x, dx')dydx \\
&= \int_A\int_C \frac{1}{h_1(x')}\int_{\mathbb{R}^n} 1_B(y)\frac{h_2(y)\Lambda(y, x)\Lambda(y', x')}{(\hat{P}_t\hat{\Lambda}_y)(x')}Q_t(\Lambda_{x'}1_D)(y)\nu_2(dy)P_t(x, dx')dx \\
&= \int_A\int_C \frac{h_1(x)}{h_1(x')}\int_{\mathbb{R}^n} \hat{Q}_t\left(1_B\frac{\hat{\Lambda}_x}{(\hat{P}_t\hat{\Lambda}_.)(x')}\right)\Lambda(y, x')1_D(y)h_2(y)dyP_t(x, dx')dx
\end{aligned}$$

$$
\begin{aligned}
&= \int_D \int_A h_1(x) \int_{\mathbb{R}^m} \hat{Q}_t \left(1_B \frac{\hat{\Lambda}_x}{(\hat{P}_t\hat{\Lambda}.)(x')} \right) 1_C(x')\hat{\Lambda}(x',y) P_t(x,dx')dxdy \\
&= \int_D \int_{\mathbb{R}^m} 1_A(x) P_t \left(1_C \hat{\Lambda}_y \hat{Q}_t \left(1_B \frac{\hat{\Lambda}_x}{(\hat{P}_t\hat{\Lambda}.)} \right) \right) \nu_1(dx)dy \qquad (4.18) \\
&= \int_D \int_C h_1(x)\hat{\Lambda}(x,y)\hat{P}_t \left(1_A \hat{Q}_t \left(1_B \frac{\hat{\Lambda}.}{(\hat{P}_t\hat{\Lambda}.)(x)} \right) \right) dxdy
\end{aligned}
$$

Lo cual muestra que, al igual que ν_2, la distribución conjunta de $(Z_0^1, Z_0^2, Z_t^1, Z_t^2)$ es absolutamente continua con respecto a la medida de Lebesgue. Así,

$$
\begin{aligned}
\lambda(x,y,A\times B) &= \frac{h_1(x)\hat{\Lambda}(x,y)\hat{P}_t \left(1_A \hat{Q}_t \left(1_B \frac{\hat{\Lambda}.}{(\hat{P}_t\hat{\Lambda}.)(x)} \right) \right)}{h_2(y)\Lambda(y,x)} \\
&= \hat{P}_t \left(1_A \hat{Q}_t \left(1_B \frac{\hat{\Lambda}.}{(\hat{P}_t\hat{\Lambda}.)(x)} \right) \right) \\
&= \int_A \hat{Q}_t \left(1_B \frac{\hat{\Lambda}_{x'}}{(\hat{P}_t\hat{\Lambda}.)(x)} \right) \hat{P}_t(x,dx') \\
&= \int_A \int_B \frac{\hat{\Lambda}(x',y')}{(\hat{P}_t\hat{\Lambda}_{y'})(x)} \hat{Q}_t(y,dy')\hat{P}_t(x,dx')
\end{aligned}
$$

como se quería. □

Ejemplos

En esta sección se comenzará exhibiendo algunos ejemplos de procesos de Markov cuyos semigrupos están entrelazados y después se procederá a mostrar algunos ejemplos de difusiones entrelazadas. Los primeros dos de ellos pueden hallarse en [1]. La siguiente proposición es la base para mostrar los entrelazamientos que se contemplan ahí:

Proposición 4.6. *Sean $(X_t)_{t\geq 0}$ y $(Y_t)_{t\geq 0}$ son dos procesos definidos en el mismo espacio de probabilidad $(\Omega, \mathscr{F}, P)$ con espacios de estados E y F respectivamente y $(\mathscr{F}_t)_{t\geq 0}$ y $(\mathscr{G}_t)_{t\geq 0}$ dos filtraciones. Supóngase que se satisfacen las siguientes cuatro condiciones:*

1. *$(\mathscr{G}_t)_{t\geq 0}$ es subfiltración de $(\mathscr{F}_t)_{t\geq 0}$.*
2. *X es adaptado a $(\mathscr{F}_t)_{t\geq 0}$ y Y es adaptado a $(\mathscr{G}_t)_{t\geq 0}$.*

3. X es de Markov respecto a $(\mathscr{F}_t)_{t\geq 0}$ con semigrupo $(P_t)_{t\geq 0}$ y Y es de Markov respecto a $(\mathscr{G}_t)_{t\geq 0}$ con semigrupo $(Q_t)_{t\geq 0}$.

4. Existe un núcleo de transición $L : E \to F$ tal que para cada función medible no negativa f se tiene

$$\mathbb{E}[f(X_t) \mid \mathscr{G}_t] = (Lf)(Y_t) \quad \text{para cada } t \geq 0$$

entonces, para cada $t, s \geq 0$

$$Q_t Lf(Y_s) = LP_t f(Y_s) \quad \text{c.s.} \tag{4.19}$$

Obsérvese que si los semigrupos son de Feller, la condición (4.19) implica entonces

$$Q_t L = LP_t \tag{4.20}$$

es decir, $Q\langle L\rangle P$.

Demostración. La ecuación 4.19 se obtiene al calcular $\mathbb{E}[f(X_{t+s}) \mid \mathscr{G}_s]$ de dos maneras distintas: por un lado, se tiene

$$\begin{aligned}
\mathbb{E}[f(X_{t+s} \mid \mathscr{G}_s] &= \mathbb{E}[\mathbb{E}[f(X_{t+s} \mid \mathscr{G}_{t+s} \mid \mathscr{G}_s] \\
&= \mathbb{E}[Lf(Y_{t+s}) \mid \mathscr{G}_s] \\
&= Q_t Lf(Y_s).
\end{aligned}$$

Y, por el otro lado, se tiene

$$\mathbb{E}[f(X_{t+s}) \mid \mathscr{G}_s] = \mathbb{E}[\mathbb{E}[f(X_{t+s} \mid \mathscr{F}_s \mid \mathscr{G}_s] = \mathbb{E}[P_t f(X_s) \mid \mathscr{G}_s] = LP_t f(Y_s)$$

□

Ejemplo 1. El primer ejemplo ha sido discutido ya después de la demostración del Teorema de Dynkin en el Capítulo 1. En este caso, Y es un movimiento browniano n-dimensional, $X_t = |Y_t|$ para todo $t \geq 0$, $\mathscr{F} = \mathscr{G}$ la filtración generada por Y y L está dada por $Lf(y) = f(|y|)$ para cualquier función medible acotada (obsérvese que se intercambió el nombre de los procesos respecto a los de aquel capítulo para obtener el entrelazamiento en el orden dado por la proposición 2). Ahora,

$$\mathbb{E}[f(X_t)g(Y_0) \mid Y_t] = Lf(Y_t)\mathbb{E}[g(Y_0) \mid Y_t] = \mathbb{E}[f(X_t) \mid Y_t]\mathbb{E}[g(Y_0) \mid Y_t]$$

y

$$\mathbb{E}[f(X_0)g(Y_t) \mid X_t, Y_0] = Lf(Y_0)\mathbb{E}[g(Y_t) \mid X_t, Y_0] = \mathbb{E}[f(X_0) \mid X_t, Y_0]\mathbb{E}[g(Y_t) \mid X_t, Y_0]$$

lo cual muestra que las condiciones (iv)' y (v) del Teorema 3.14 son satisfechas. Se afirma que, sin embargo, L no admite una densidad Λ. De ser así, se podría tomar$f = 1_{\{0\}}$, por lo que

$$1 = f(|0|) = Lf(0) = \int_{\mathbb{R}} \Lambda(0,x)f(x)dx = \int_{\{0\}} \Lambda(0,x)dx = 0$$

lo cual claramente no es posible. Así, no todas las condiciones del Teorema 3.14 son satisfechas. Usando que la ecuación diferencial estocástica para X es

$$dX_t = \frac{n-1}{2X_t}dt + 2\sum_{i=1}^{n} \frac{Y_t^i}{X_t} dY_t^i$$

entonces, para cualquier $f \in C_0(\mathbb{R}^{n+1})$ y estado inicial $z = (x, y)$

$$\mathbb{E}_z[f(X_t, Y_t)] = f(z) + \mathbb{E}_z\Big[\frac{n-1}{2}\int_0^t \frac{\partial f}{\partial x_{n+1}}(X_s, Y_s)\frac{1}{X_s} + \frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2 f}{\partial x_i^2}(X_s, Y_s)$$
$$+ 2\sum_{i=1}^{n}\frac{\partial^2 f}{\partial x_{n+1}\partial x_i}(X_s, Y_s)\frac{Y_s^i}{X_s} + \frac{\partial^2 f}{\partial x_{n+1}^2}(X_s, Y_s)X_s ds\Big]$$

por lo que el generador de (X, Y) tiene términos cruzados de la primer variable con todas las demás, por lo que no es de la forma enunciada en el Teorema 3.3.

Ejemplo 2. Para $\alpha, \beta > 0$, sean X^α y X^β los cuadrados de dos procesos de Bessel independientes de dimensiones 2α y 2β, respectivamente, los cuales inician en 0. Sea $X = X^\alpha$ y $Y = X^\alpha + X^\beta$ (por lo que Y es también el cuadrado de un proceso de Bessel de dimensión $2(\alpha + \beta)$) y sean $\mathscr{F}$ y $\mathscr{G}$ las filtraciones generadas por el par (X, Y) y Y respectivamente. En las funciones medibles acotadas en $[0, \infty)$, defínase el operador de transición estocástico $L_{\alpha,\beta}$ por

$$L_{\alpha,\beta}f(y) = \frac{1}{B(\alpha, \beta)}\int_0^1 f(yz)z^{\alpha-1}(1-z)^{\beta-1}dz$$

donde B es la función beta. Después de un cambio de variable, se puede ver que la densidad de este operador está dada por

$$\Lambda_{\alpha,\beta}(y, x) = \frac{1}{yB(a, b)}\Big(\frac{x}{y}\Big)^{\alpha-1}\Big(1-\frac{x}{y}\Big)^{\beta-1}1_{(0,y)}(x) \tag{4.21}$$

Si los semigrupos $(Q_t)_{t\geq 0}$ y $(P_t)_{t\geq 0}$ son los asociados a X y Y respectivamente, se tiene la siguiente

Proposición 4.7. *Para cada $t \geq 0$, $Q_t L_{\alpha,\beta} = L_{\alpha,\beta}P_t$, es decir, $Q\langle L_{\alpha,\beta}\rangle P$.*

Demostración. La identidad se obtendrá como consecuencia de la proposición . Por inversión temporal, los procesos $(t^2 X^\alpha_{1/t})_{t\geq 0}$ y $(t^2 X^\beta_{1/t})_{t\geq 0}$ son cuadrados de procesos de Bessel independientes de dimensiones 2α y 2β, respectivamente.
Sea H una funcional medible no negativa y sea f una función Borel medible no negativa. Así,

$$\mathbb{E}[H(Y_u, u \leq t)f(X_t^\alpha)] = \mathbb{E}[H(u^2 Y_{1/u}, u \leq t)f(t^2 X_{1/t})] \tag{4.22}$$

Sea $H_t = H(u^2 Y_{1/u}, u \leq t)$. Puesto que $(t^2 Y_{1/t})_{t\geq 0}$ es de Markov respecto a $\mathscr{G}$,

$$\mathbb{E}[H(Y_u, u \leq t) f(X_t^\alpha)] = \mathbb{E}[H_t f(t^2 X_{1/t}^\alpha)] = \mathbb{E}[\mathbb{E}[H_t \mid Y_{1/t}] f(t^2 X_{1/t}].$$

Usando que para dos variables aleatorias independientes $U_{\alpha,\beta}$ y $U_{\alpha+\beta}$ con distribución $Beta(\alpha, \beta)$ y $Gamma(\alpha + \beta, 1)$ respectivamente y dos variables aleatorias independientes U_α y U_β con distribución $Gamma(\alpha, 1)$ y $Gamma(\beta, 1)$ se tiene

$$(U_{\alpha,\beta}, U_{\alpha+\beta}) \stackrel{\mathrm{d}}{=} \Big(\frac{U_\alpha}{U_\alpha + U_\beta}, U_\alpha + U_\beta\Big)$$

entonces, manteniendo la notación y notando que $L_{\alpha,\beta} f(x) = \mathbb{E}[f(xU_{a,b}]$,

$$\begin{aligned}
\mathbb{E}[H(Y_u, u \leq t) f(X_t^\alpha)] &= \mathbb{E}\Big[\mathbb{E}[H_t \mid Y_{1/t}] f\Big(t^2 \frac{X_{1/t}}{Y_{1/t}} Y_{1/t}\Big)\Big] \\
&= \mathbb{E}[\mathbb{E}[H_t \mid Y_{1/t}] f(t^2 U_{\alpha,\beta} Y_{1/t})] \\
&= \mathbb{E}[H_t L_{\alpha,\beta} f(t^2 Y_{1/t})] \\
&= \mathbb{E}[H(Y_u, u \leq t) L_{\alpha,\beta} f(Y_t)]
\end{aligned} \tag{4.23}$$

Comparando las ecuaciones (4.22) y (4.23) y observando que H es arbitraria, se obtiene

$$\mathbb{E}[f(X_t^\alpha) \mid \mathscr{G}_t] = L_{\alpha,\beta} f(Y_t). \tag{4.24}$$

□

Obsérvese que la ecuación (4.24) verifica la condición (iv) de la Definición 3.2 y, en particular, se satisface la condición (iv)' del Teorema 3.14. La condición (v) del mismo teorema se satisface por ser los estados iniciales constantes. Sin embargo, $\Lambda_{\alpha,\beta}$ no satisface las condiciones de regularidad impuestas (pues no es estrictamente positiva), por lo que el Teorema 3.14 no puede ser aplicado. Como en el ejemplo anterior, el generador de la difusión (X, Y) contiene términos cruzados, por lo que no tiene la forma del Teorema 3.3.

Ejemplo 3. Supóngase que los espacios de estados $\mathcal{X}$ y $\mathcal{Y}$ de las difusiones X y Y son compactos y que X tiene una distribución estacionaria μ en $\mathcal{X}$ con densidad continua positiva h. Así, para cualquier función continua $f \in \mathcal{D}(\mathcal{A}^X)$,

$$\begin{aligned}
\int_{\mathcal{X}} (\mathcal{A}^X f)(x) h(x) dx &= \int_{\mathcal{X}} (\mathcal{A}^X f) \mu(dx) \\
&= \lim_{t\to 0} \frac{1}{t} \int_{\mathcal{X}} (P_t f)(x) - f(x) \mu(dx) = 0
\end{aligned}$$

donde la tercer igualdad se debe al Teorema de convergencia dominada (por ser f acotada) y la última igualdad a que μ es una distribución estacionaria (por lo que ambas integrales $\int_{\mathcal{X}} (P_t f)(x) \mu(dx)$

y $\int_{\mathcal{X}} f(x)\mu(dx)$ calculan $\mathbb{E}_\mu[f(X_0)]$). Así, si u es una función continua en $\mathcal{X}\times\mathcal{Y}$ que satisface (3.3), para cualquier constante M, la función $u+Mf$ es también una solución de (3.3) y, en particular, para M suficientemente grande se obtienen soluciones positivas. Por lo tanto, las funciones $u+Mf$ proveen entrelazamientos de acuerdo al Teorema 4.1.
Podría preguntarse cómo afecta la elección de M al entrelazamiento. Asumiendo que $\tau(y) = \int_{\mathcal{X}} u(x,y)dx$ es continuamente diferenciable en y, el generador de la transformada h de Doob de Y asociado con $u+Mf$ según el Teorema 4.1 y el Lema 3.4 es

$$\mathcal{A}^{\tau,M} = \mathcal{A}^Y + (\nabla \log(\tau+M))'\rho\nabla_y = \mathcal{A}^Y + \frac{(\nabla\tau)'}{\tau+M}\rho\nabla_y$$

(obsérvese que $\mathcal{A}^Y\tau = 0$, pues debido a la compacidad, la prueba del Teorema 4.1 muestra que $\tau(Y)$ es en realidad es una martingala). Si adicionalmente la tercia $(\mathcal{A}^X, \mathcal{A}^{\tau,M}, u+Mf)$ satisface las hipótesis del Teorema 3.3, el generador del entrelazamiento correspondiente estaá dado por

$$\mathcal{A}^X + \mathcal{A}^Y + \Big(\frac{(\nabla\tau)'}{\tau+M} + \frac{(\nabla_y u)'}{u+Mf}\Big)\rho\nabla_y$$

lo cual muestra cómo las diferentes elecciones de M afecta a las difusiones que involucra.
Para ejemplificar esta construcción considérese $\mathcal{X} = \mathcal{Y} = [-1,1]$ y sea

$$\mathcal{A}^X = -2x\frac{d}{dx} + (1-x^2)\frac{d^2}{dx^2}, \quad \mathcal{A}^Y = (1-2y)\frac{d}{dy} + (1-y^2)\frac{d^2}{dy^2}.$$

Si f y g son dos funciones C^2, entonces

$$\begin{aligned}
\int_{-1}^{1}(\mathcal{A}^X f)(x)g(x)dx &= \int_{-1}^{1}(-2xf'(x) + (1-x^2)f''(x))g(x)dx \\
&= (1-x^2)f'(x)g(x)\mid_{-1}^{1} - \int_{-1}^{1}(1-x^2)f'(x)g'(x)dx \\
&= -(1-x^2)f(x)g'(x)\mid_{-1}^{1} + \int_{-1}^{1} f(x)(-2xg'(x) + (1-x^2)g''(x))dx \\
&= \int_{-1}^{1} f(x)(\mathcal{A}^X g)(x)dx
\end{aligned}$$

por lo que $\mathcal{A}^X$ coincide con $(\mathcal{A}^X)^*$ en las funciones C^2 (en el sentido de que, si μ es la medida inducida por $f \in \mathcal{D}(\mathcal{A}^X) \cap C^2([-1,1])$, entonces $(\mathcal{A}^X)^*\mu$ tiene densidad $\mathcal{A}^X f$). El operador $\mathcal{A}^X$ tiene eigenfunciones $(f_n)_{n\in\mathbb{N}}$ tal que el eigenvalor asociado a f_n es $n(n+1)$ y, del mismo modo, el operador $\mathcal{A}^Y$ tiene eigenfunciones $(g_n)_{n\in\mathbb{N}}$ de manera que el eigenvalores de g_n también es $n(n+1)$, (consúltese, por ejemplo, [10], página 73, Teorema 3.16). Por lo tanto, las funciones de la forma $u(x,y) = \sum_{n=1}^{\infty} c_n f_n(x)g_n(y)$ satisfacen (3.3) siempre y cuando $\sum_{n=1}^{\infty}|c_n|\|f_n\|_\infty\|g_n\|_\infty < \infty$ y

$\sum_{n=1}^{\infty} n(n+1)|c_n|\|f_n\|_\infty\|g_n\|_\infty < \infty$.
Para finalizar el ejemplo, hace falta hallar una distribución estacionaria para X con densidad en $[-1,1]$. Se afirma que la distibución uniforme ν en dicho intervalo satisface esto. Si se denota por h la densidad de ν, entonces h es constante, por lo que $\mathcal{A}^X h = 0$. Así, por lo demostrado anteriormente, para cualquier función f en $C_0^\infty([-1,1])$,

$$\int_{-1}^{1} \mathcal{A}^X g(x)\nu(dx) = \int_{-1}^{1} g(x)(\mathcal{A}^X h)(x)dx = 0$$

y, por lo tanto,

$$\begin{aligned}
\mathbb{E}_\nu[f(X_t)] &= \int_{-1}^{1} \mathbb{E}_x[f(X_t)\nu(dx) \\
&= \int_{-1}^{1} f(x) + \mathbb{E}_x\Big[\int_0^t \mathcal{A}^X f(X_s)ds\Big]\nu(dx) \\
&= \int_{-1}^{1} f(x)\nu(dx) + \int_{-1}^{1}\int_0^t P_s\mathcal{A}^X f(x)ds\nu(dx) \\
&= \mathbb{E}_\nu[f(X_0)] + \int_0^t\int_{-1}^{1} \mathcal{A}^X P_s f(x)\nu(dx)ds = \mathbb{E}_\nu[f(X_0)]
\end{aligned}$$

lo cual demuestra que ν en efecto es una distribución estacionaria. Luego, las funciones de la forma $\frac{M}{2}+\sum_{n=1}^{\infty} c_n f_n(x)g_n(y)$ satisfacen (3.3) y son positivas para todo $M > 2\sum_{n=1}^{\infty}\sum_{n=1}^{\infty}|c_n|\|f_n\|_\infty\|g_n\|_\infty$, dando origen a entrelazamientos de X y transformadas h de Doob de Y.

Ejemplo 4. Para exponer el siguiente ejemplo, se utilizará el siguiente lema

Lema 4.8. *Para $m > 1$, sea u una función de densidad en $\mathbb{R}^m$ dos veces diferenciable y estrictamente positiva. Sea $\gamma_m = \pi^{m/2}/\Gamma(1+m/2)$ el volumen de la esfera unitaria m-dimensional. Para $r > 0$, defínase*

$$v(r,x) = \frac{1}{m\gamma_m}\int_{\partial B(0,1)} u(x+rz)\theta(dz) \tag{4.25}$$

donde $B(0,1)$ es la bola unitaria centrada en 0 y θ es la medida de Lebesgue en su frontera (por lo que la integral es solo una integral de superficie). Entonces, v es una solución positiva de

$$\frac{m-1}{2r}\frac{\partial v}{\partial r} + \frac{1}{2}\frac{\partial^2 v}{\partial r^2} = \frac{1}{2}\Delta_x v \tag{4.26}$$

Demostración. La prueba está basada principalmente en las ecuaciones obtenidas en la demostración del Lema 1 de la página 71 en [3]. Allí, se demuestra que

$$\frac{\partial v}{\partial r}(x,r) = \frac{r}{m^2\gamma_m}\int_{B(0,1)} \Delta u(x+rz)dz \quad \text{y}$$

$$\frac{\partial^2 v}{\partial r^2}(x,r) = \frac{1}{m\gamma_m}\int_{\partial B(0,1)} \Delta u(x+rz)\theta(dz) + \frac{m-1}{m^2\gamma_m}\int_{B(0,1)} \Delta u(x+rz)dz$$

de donde se sigue el lema, ya que $\Delta_x v(r,x) = \frac{1}{m\gamma_m}\int_{\partial B(0,1)} \Delta u(x+rz)\theta(dz)$ $\square$

Por el Teorema de Fubini,

$$\int_{\mathbb{R}^m} v(r,x)dx = \frac{1}{m\gamma_m}\int_{\partial B(0,1)}\int_{\mathbb{R}^m} u(x+rz)dx\theta(dz) = \frac{1}{m\gamma_m}\int_{\partial B(0,1)} \theta(dz) = 1$$

por lo que el operador inducido por v es estocástico. Esto permite usar el Teorema 3.3 para construir entrelazamientos de movimientos brownianos multidimensionales con procesos de Bessel de la misma dimensión. Obsérvese que estos entrelazamientos son distintos a los provistos por el ejemplo 1.

De manera más general, el Teorema 4.1 implica que cualquier solución positiva de (4.26) da origen a entrelazamientos de movimientos brownianos multidimensionales con transformadas h de procesos de Bessel de la misma dimensión. Las posibles transformadas h de dichos procesos son caracterizadas por la siguiente proposición

Proposición 4.9. *Sea v una solución positiva de (4.26) con $m > 1$. Supóngase que $\int_{\mathbb{R}^m} |\Delta_x u(r,x)|dx$ está localmente acotada conforme r varía y que la integral $\tau(r) = \int_{\mathbb{R}^m} u(r,x)dx$ es finita para todo $r > 0$ y continua en r. Entonces, existen constantes $a, b \in \mathbb{R}$ tales que $\tau(t) = a + br^{2-m}$ si $m > 2$ y $\tau(r) = a + b\log r$ si $m = 2$. En particular, si $\limsup_{\tau\downarrow 0} |\tau(r)| < \infty$, entonces τ es constante.*

Demostración. Denótese por Y al proceso de Bessel. Según la observación hecha después del enunciado del Teorema 4.1, las condiciones en v permiten aplicar dicho Teorema para concluir que $\tau(Y)$ es una martingala local. Por las proposiciones 3.2 y 3.5 del Capítulo VII de [7], τ es una función de escala para Y y, salvo una transformación afín, existe solo una de ellas. Finalmente, en el segundo párrafo de la página 446 en [7] se da una expresión para las funciones escala de Y, donde se ve que si $m > 2$, la función se puede tomar como $y \mapsto -y^{2-m}$ y si $m = 2$, se puede tomar $y \mapsto 2\log y$ en su lugar. $\square$

BIBLIOGRAFÍA

[1] P. Carmona, F. Petit, and M. Yor. Beta-gamma random variables and intertwining relations between certain markov processes. *Rev. Mat. Iberoamericana*, 14, 1998.

[2] K. Engel, S. Brendle, R. Nagel, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, et al. *One-Parameter Semigroups for Linear Evolution Equations*. Graduate Texts in Mathematics. Springer New York, 2006.

[3] L. C. Evans. *Partial differential equations*. American Mathematical Society, 2010.

[4] O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer New York, 2010.

[5] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.

[6] S. Pal and M. Shkolnikov. Intertwining diffusions and wave equations. *arXiv:1306.0857*, 2013.

[7] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 2013.

[8] L. C. G. Rogers and J. W. Pitman. Markov functions. *Ann. Probab.*, 9(4):573–582, 08 1981.

[9] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes and martingales: Volume 2, Itô calculus*, volume 2. Cambridge university press, 2000.

[10] J. Shen, T. Tang, and L.-L. Wang. *Spectral methods: algorithms, analysis and applications*, volume 41. Springer Science & Business Media, 2011.