



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE CIENCIAS**

**ANÁLISIS ESPECTRAL DE PROCESOS DE  
DIFUSIÓN CAMBIANTES**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE:**

**Matemático Aplicado**

**PRESENTA :**

**Ayrton Pablo Almada Jiménez**

**DIRECTOR DE TESIS:**

**Manuel Domínguez de la Iglesia**

**Ciudad Universitaria, CDMX.**

**Mayo 2021.**





Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# Contenido

	Página
<b>Introducción</b>	<b>3</b>
<b>1 Capítulo I: Preliminares</b>	<b>5</b>
1.1 Elementos de integración . . . . .	5
1.2 Procesos de difusión . . . . .	6
1.3 Polinomios ortogonales y medidas matriciales . . . . .	12
1.4 Resultados adicionales . . . . .	17
<b>2 Capítulo II: Construcción y desarrollo de procesos de difusión cambiante</b>	<b>24</b>
2.1 Procesos de difusión cambiante y propiedades importantes . . . . .	24
2.2 Regularidad de un proceso de difusión cambiante . . . . .	36
2.3 Continuidad de un proceso de difusión cambiante . . . . .	39
2.4 Recurrencia . . . . .	52
2.5 Ergodicidad . . . . .	75
<b>3 Capítulo III: Análisis espectral de procesos de difusión cambiante</b>	<b>86</b>
3.1 El operador infinitesimal . . . . .	86
3.2 Ecuaciones matriciales de retroceso y de evolución del proceso . . . . .	87
3.3 Representación espectral de la función de densidad de transición matricial . . . . .	89
3.4 Ecuaciones diferenciales asociadas a algunos operadores . . . . .	90
3.5 Ejemplo de análisis espectral de procesos de difusión cambiante . . . . .	93
<b>Referencias</b>	<b>106</b>

# Introducción

Desde 1930 matemáticos como N. Wiener o K. Itô ya sabían de la conexión entre polinomios ortogonales y procesos estocásticos, en particular la relación entre los polinomios de Hermite y la teoría de integración con respecto al movimiento Browniano. Más adelante, a partir de 1950, autores como M. Kac, E. Hille, W. Feller, H.P. McKean, S. Karlin, J. McGregor, J.F. Barrett y D.G. Lampard, entre otros, establecen una serie de conexiones importantes entre las funciones de probabilidad de transición de procesos de difusión, procesos de nacimiento y muerte a tiempo continuo y cadenas de nacimiento y muerte a tiempo discreto (en ese orden) todo en términos de una **representación espectral**. Esta representación espectral está basada en el análisis espectral del operador infinitesimal asociado con este tipo de procesos de Markov y otros aspectos probabilísticos que pueden analizarse en términos de las correspondientes eigenfunciones y eigenvalores.

El objetivo principal de esta tesis es definir los conceptos fundamentales para desarrollar una metodología y realizar el análisis espectral de los denominados procesos de difusión cambiante, así como analizar las características esenciales de este tipo de procesos. El concepto de proceso de difusión cambiante se refiere a la coexistencia de un proceso a tiempo continuo y de espacio de estados real, denominado nivel, con un proceso a tiempo continuo y de espacio de estados discreto, denominado fase. Esta clase de procesos suelen describirse en términos de operadores infinitesimales asociados a sistemas de ecuaciones diferenciales estocásticas. A todo esto se le añaden las leyes de transición de probabilidad del proceso entre las correspondientes fases. La importancia de estos conceptos recae en las aplicaciones recientes y en los constantes avances en comunicaciones inalámbricas, procesamiento de señales, sistemas biológicos, ingeniería financiera, redes de cola, modelación, análisis, control y optimización de sistemas a gran escala, todo bajo la influencia de entornos aleatorios.

Para llevar a cabo el análisis espectral, nos vamos a enfocar principalmente en los procesos de difusión cambiante cuyo nivel tenga espacio de estados unidimensional y un número finito de fases. Esto debido a la dificultad en el desarrollo del análisis que implica el aumentar la dimensión en el espacio de estados, (eso sin mencionar la dificultad computacional que implica la simulación de procesos multidimensionales). De manera similar a la mayoría de los textos que desarrollan la teoría espectral de procesos estocásticos, el desarrollo de la teoría en esta tesis será muy técnica. En este trabajo también nos enfocaremos en la parte práctica al simular estos procesos, todo esto apoyándonos en el lenguaje de programación **R** para, por un lado, resolver las ecuaciones diferenciales estocásticas por medio de aproximaciones numéricas y, por otro lado, generar gráficas que ayudarán a ilustrar los ejemplos.

El primer capítulo consta de definiciones preliminares de teoría de probabilidad e integración, procesos estocásticos, polinomios ortogonales y medidas matriciales, puesto que todo este material sustenta la construcción de los procesos de difusión cambiante y el análisis espectral de operadores infinitesimales. La fuente principal de esta parte es [36].

En el segundo capítulo se lleva a cabo la construcción y desarrollo de procesos de difusión cambiante. Primero se describen formalmente las componentes de esta clase de procesos, es decir trabajaremos con procesos de la forma  $(X(t), Y(t))$  que satisfacen ecuaciones diferenciales estocásticas de la forma:

$$\begin{aligned}dX(t) &= b(X(t), Y(t))dt + \sigma(X(t), Y(t))dw(t), \\(X(0), Y(0)) &= (x, \alpha) : x \in S \subset \mathbb{R}, \alpha \in \{1, 2, \dots, N\},\end{aligned}$$

donde  $w(t)$  denota el movimiento Browniano regular. Después se describe y desarrolla una serie de conceptos importantes: regularidad, continuidad del proceso, propiedad de Feller débil y fuerte, recurrencia, transitividad, recurrencia nula/positiva y ergodicidad. Las fuentes principales de esta parte son [36] y [4].

En el tercer y último capítulo se lleva a cabo el análisis espectral de procesos de difusión cambiante, es decir, dado el proceso  $(X(t), Y(t))$  previamente mencionado buscamos una serie de eigenfunciones y eigenvalores con los cuales se pueda expresar la función de transición de probabilidad del proceso como una serie cuyos componentes estén determinados por esos eigen pares. Como aplicación se podrá obtener a través de estos términos las distribuciones estacionarias e invariantes del proceso, si existiesen. También se estudiará en

detalle un ejemplo que es una variante del modelo de difusión de Wright-Fisher que involucra solo efectos de mutación. La fuente principal de esta parte es [12].

# 1 Capítulo I: Preliminares

En este capítulo proporcionaremos definiciones básicas del concepto de *medida* para definir la integral con la cual daremos sentido a la *ortogonalidad de polinomios de variable real*. Seguido se define el concepto de *proceso de difusión* junto con algunas definiciones y propiedades con respecto a esta clase de procesos. Posteriormente extenderemos el concepto de *ortogonalidad para polinomios matriciales* para luego definir el concepto de *medida matricial* y enunciar algunas propiedades. Finalmente enunciamos una serie de definiciones y resultados que, más adelante, en los Capítulos II y III, serán de gran utilidad para demostrar varios lemas, proposiciones y teoremas.

## 1.1 Elementos de integración

**Definición 1.1.1:** Sea  $X \neq \emptyset$  un conjunto. Una familia  $\mathcal{X}$  de subconjuntos de  $X$  es una  $\sigma$ -álgebra si cumple:

- $\emptyset, X \in \mathcal{X}$ .
- Si  $A \in \mathcal{X}$ , entonces el complemento  $A^C \in \mathcal{X}$ .
- $\{A_n\}_{n \in \mathbb{N}} \in \mathcal{X} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{X}$ .

Un ejemplo de lo anteriormente mencionado es la  $\sigma$ -álgebra de Borel  $\mathbb{B}$ , generada por todos los intervalos abiertos  $(a, b) \in \mathbb{R}$ .

Denotemos por  $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$  al conjunto de los *reales extendidos*.

**Definición 1.1.2:** Una familia de  $\sigma$ -álgebras,  $\{\mathcal{F}_t\}$ , para todo  $t \geq 0$  o  $t = 1, 2, \dots$ , se le llama *filtración*, si para cualesquiera  $s \leq t$  se tiene que  $\mathcal{F}_s \subset \mathcal{F}_t$ . Decimos que  $\mathcal{F}_t$  es *completo* si contiene cualquier conjunto nulo y también diremos que  $\{\mathcal{F}_t\}$  *satisface la condición usual* si  $\mathcal{F}_0$  es completo.

A un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$  dotado de una filtración  $\{\mathcal{F}_t\}$  se le llama *espacio de probabilidad filtrado* y lo denotamos por  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ .

**Definición 1.1.3:** Sea  $f : X \rightarrow \overline{\mathbb{R}}$ , decimos que  $f$  es  $\mathcal{X}$ -medible si se satisface lo siguiente:

$$\forall a \in \mathbb{R} \quad \{x \in X : f(x) > a\} \in \mathcal{X}.$$

Denotamos por  $M(X, \mathcal{X})$  al conjunto de las funciones medibles con valores en los reales extendidos.

**Definición 1.1.4:** Una medida de la  $\sigma$ -álgebra  $\mathcal{X}$  es una función  $\mu : \mathcal{X} \rightarrow \overline{\mathbb{R}}_+$  que satisface las siguientes condiciones:

- $\mu(\emptyset) = 0$ .
- Sea  $\{E_n\}_{n \in \mathbb{N}}$  cualquier sucesión disjunta de subconjuntos de  $\mathcal{X}$ , entonces:

$$\mu \left( \bigcup_{n=1}^{\infty} E_n \right) = \sum_{n=1}^{\infty} \mu(E_n).$$

Si ajustamos  $X = \mathbb{R}, \mathcal{X} = \mathbb{B}$ , donde  $\mathbb{B}$  es la  $\sigma$ -álgebra de Borel, entonces existe una única medida  $\lambda$  que coincide con la medida de los intervalos (es decir,  $\forall (a, b) \in \mathcal{X} \quad \lambda((a, b)) = b - a$ ). A esta medida  $\lambda$  le llamamos la *medida de Lebesgue*.

En el caso de que una medida  $\mu$  no tome el valor  $+\infty$ , decimos que la medida es *finita*. Por otro lado, una medida es  $\sigma$ -finita si existe una sucesión de subconjuntos  $\{E_n\}_{n \in \mathbb{N}}$  de  $\mathcal{X}$  tal que  $\mathcal{X} = \bigcup_{n=1}^{\infty} E_n$  y

$\forall n \in \mathbb{N} \quad \mu(E_n) < +\infty$ . Nótese que la medida de Lebesgue  $\lambda$  también es  $\sigma$ -finita, pues  $\mathbb{R} = \bigcup_{k \in \mathbb{Z}} \{[k, k+1)\}$  y  $\forall k \in \mathbb{Z} \quad \mu([k, k+1)) = 1 < +\infty$ .

Ahora que tenemos definido el espacio de medida  $(X, \mathcal{X}, \mu)$ , vamos a definir la integral con respecto a la medida  $\mu$ . Primero para el caso de funciones simples, luego funciones no negativas y finalmente se define el concepto de funciones integrables, todo esto antes de pasar a la sección de polinomios ortogonales.

**Definición 1.1.5:** Una función medible y de variable real es simple si y solo si tiene un número finito de valores. Una función simple medible es de la forma:

$$\phi(x) = \sum_{j=1}^n a_j \mathbb{1}_{E_j}(x),$$

donde  $a_j \in \mathbb{R}$  y  $\mathbb{1}_{E_j}$  es la función indicadora del conjunto  $E_j \in \mathcal{X}$ .

**Definición 1.1.6:** Sea  $\phi(x) = \sum_{j=1}^n a_j \mathbb{1}_{E_j}(x)$  función simple y medible, definimos la integral de  $\phi$  con respecto a  $\mu$  como:

$$\int_X \phi d\mu = \sum_{j=1}^n a_j \mu(E_j).$$

**Definición 1.1.7:** Sea  $f \in M^+(X, \mathcal{X})$  una función medible y no negativa, definimos la integral de  $f$  con respecto a  $\mu$  como:

$$\int_X f d\mu = \sup_{\phi \in F} \int_X \phi d\mu,$$

donde  $F = \{\phi \text{ función simple y medible} : 0 \leq \phi(x) \leq f(x) \quad \forall x \in X\}$ .

**Definición 1.1.8:** Sea  $f \in M(X, \mathcal{X})$  una función medible, decimos que es integrable si sus partes negativa y positiva  $f^+, f^-$ , tienen integrales finitas con respecto a  $\mu$ , de tal forma que la integral de  $f$  con respecto a  $\mu$  se define como:

$$\int_X f d\mu = \int_X f^+ d\mu + \int_X f^- d\mu.$$

## 1.2 Procesos de difusión

En esta sección ocuparemos como referencias principales [36] y [6].

**Definición 1.2.1:** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $\{X_t : t \geq 0\}$  un proceso de Markov a tiempo continuo cuyo espacio de estados es un intervalo real  $\mathbb{S} = (a, b) \subset \mathbb{R}, -\infty \leq a < b \leq +\infty$  y  $p(s, x, t, y)$  la función de densidad de transición del proceso, decimos que  $\{X_t : t \geq 0\}$  es un proceso de difusión si existen funciones  $b(t, x)$ ,  $\sigma^2(t, x)$  conocidas como coeficiente de deriva y de difusión respectivamente, las cuales satisfacen que para todo  $\epsilon > 0$ ,  $x, y \in \mathbb{S}$ ,  $t, \Delta t \geq 0$ :

$$\begin{aligned} \int_{|y-x|=\epsilon} p(t, x, t + \Delta t, y) dy &= o(\Delta t), \\ \int_{|y-x| \leq \epsilon} (y-x) p(t, x, t + \Delta t, y) dy &= b(t, x) + o(\Delta t), \\ \int_{|y-x| \leq \epsilon} (y-x)^2 p(t, x, t + \Delta t, y) dy &= \sigma^2(t, x) + o(\Delta t), \end{aligned} \tag{1}$$

donde  $o(\Delta t)$  satisface que

$$\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0.$$

Se suele suponer que la convergencia de estas relaciones es uniforme con respecto a la variable temporal  $t$  dentro de todo intervalo finito  $t_0 \leq t \leq t_1$  y también con respecto a  $x$ .

**Definición 1.2.2:** Ahora, supóngase que  $\{X_t : t \geq 0\}$  es homogéneo con respecto al tiempo, posee trayectorias continuas y no necesariamente tiene incrementos independientes (tal como el caso del movimiento Browniano). Asumiendo que  $X_s = x$  y para valores pequeños de  $t$ , el incremento  $X_{t+s} - X_s = X_{t+s} - x$  tiene media y varianza de valor aproximado  $tb(x)$  y  $t\sigma^2(x)$  respectivamente, (nótese que dado el supuesto de homogeneidad temporal, podemos asumir que el coeficiente de deriva y de difusión no dependen del tiempo). Entonces diremos que  $\{X_t : t \geq 0\}$  es un proceso de difusión si satisface:

$$\begin{cases} \mathbb{E}[X_{t+s} - X_s | X_s = x] = tb(x) + o(t), \\ \mathbb{E}[(X_{t+s} - X_s)^2 | X_s = x] = t\sigma^2(x) + o(t), \\ \mathbb{E}[|X_{t+s} - X_s|^3 | X_s = x] = o(t). \end{cases} \quad (2)$$

Por lo tanto un proceso de Markov  $\{X_t : t \geq 0\}$  con espacio de estados  $\mathbb{S}$  es un proceso de difusión con coeficiente de deriva  $b(x)$  y coeficiente de difusión  $\sigma^2(x)$  si posee trayectorias continuas y se satisfacen las anteriores relaciones. En caso de que los coeficientes de deriva y difusión sean constantes nos referiremos a  $\{X_t : t \geq 0\}$  como un *Movimiento Browniano con deriva y reescalamiento*.

**Definición 1.2.3:** Definimos la función de distribución de probabilidad del proceso  $\{X_t : t \geq 0\}$  como  $p(t; x, dy)$  dada por:

$$p(t; x, dy) = \mathbb{P}_x[X_t \in dy] = \mathbb{P}[X_t \in dy | X_0 = x].$$

Si  $p(t; x, dy)$  posee una función de densidad  $p(t; x, y)$ , entonces para todo conjunto de Borel  $B \subset \mathbb{S}$  tendremos que:

$$p(t; x, B) = \int_B p(t; x, y) dy.$$

Esta función existe si  $b(x), \sigma(x) \in \mathcal{C}^1(\mathbb{S})$ , es decir funciones continuamente diferenciables con derivadas acotadas en  $\mathbb{S}$ , además  $\sigma''(x)$  existe, es continua y  $\sigma^2(x) > 0$  para todo  $x \in \mathbb{S}$ .

Considérese el conjunto de todas las funciones acotadas y Borel medibles cuyo dominio sea el conjunto  $\mathbb{S}$ , lo denotaremos por  $\mathbb{B}(\mathbb{S})$ . Sobre este espacio definimos al *operador de transición del proceso*  $\{X_t : t \geq 0\}$  de la siguiente manera:

**Definición 1.2.4:** El operador de transición del proceso  $\{X_t : t \geq 0\}$  aplicado a  $f \in \mathbb{B}(\mathbb{S})$ , es el siguiente:

$$(T_t f)(x) = \mathbb{E}_x[f(X_t)] := \mathbb{E}[f(X_t) | X_0 = x] = \int_{\mathbb{S}} f(y) p(t; x, dy), \text{ donde } t > 0, x \in \mathbb{S}. \quad (3)$$

Este operador sustituye al concepto de matriz de transición de probabilidades para cadenas de nacimiento y muerte a tiempo discreto o funciones de probabilidad de transición para procesos de nacimiento y muerte.

Por ejemplo, si  $f(x) = \mathbb{1}_B(x)$ , (i.e.,  $f$  es la función indicadora sobre el conjunto  $B \subset \mathbb{S}$ ), tendremos que:

$$(T_t \mathbb{1}_B)(x) = \mathbb{P}[X_t \in B | X_0 = x].$$



El operador de transición  $T_t : \mathbb{B}(\mathbb{S}) \mapsto \mathbb{B}(\mathbb{S})$  es un operador lineal y acotado sobre el conjunto  $\mathbb{B}(\mathbb{S})$  bajo la norma del supremo  $\|f\| = \sup\{|f(y)| : y \in \mathbb{S}\}$  y  $(T_0f)(x) = f(x)$  para toda  $f \in \mathbb{B}(\mathbb{S})$ . De hecho,  $T_t$  es una *contracción*, i.e.  $\|T_t\| \leq \|f\|$  para toda  $f \in \mathbb{B}(\mathbb{S})$ , eso debido a que:

$$|(T_t f)(x)| \leq \int_{\mathbb{S}} |f(y)| p(t; x, dy) \leq \|f\|.$$

La familia de operadores de transición  $\{T_t : t > 0\}$  posee la *propiedad de semigrupo*  $T_{s+t} = T_s T_t$ . De hecho, ocupando la conocida propiedad  $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$ , se tiene que:

$$(T_{s+t}f)(x) = \mathbb{E}[f(X_{s+t})|X_0 = x] = \mathbb{E}[\mathbb{E}[f(X_{s+t})|X_s]|X_0 = x] = \mathbb{E}[(T_t f)(X_s)|X_0 = x] = T_s(T_t f)(x). \quad (4)$$

En consecuencia tenemos que  $T_s T_t = T_t T_s$ . Esta propiedad de semigrupo implica que el comportamiento de  $T_t$  está completamente determinado por el comportamiento de  $T_t$  cuando  $t \downarrow 0$  (muy similar al caso de las cadenas de Markov a tiempo continuo).

**Definición 1.2.5:** Sea  $\{T_t : t > 0\}$  el operador de transición (3) de un proceso de difusión  $\{X_t : t > 0\}$ . Definimos el operador de infinitesimal  $\mathcal{L}$  de  $\{T_t : t > 0\}$  como el operador lineal de la forma

$$(\mathcal{L}f)(x) = \lim_{t \downarrow 0} \frac{(T_t f)(x) - f(x)}{t},$$

para toda  $f \in \mathbb{B}(\mathbb{S})$  de forma tal que el lado derecho converga uniformemente en  $x$  a cierta función. La clase de todas estas funciones  $f$  forma el dominio  $D_{\mathcal{L}}$  de  $\mathcal{L}$ .

**Proposición 1.2.6:** Sea  $f \in \mathcal{C}^2(\mathbb{S})$  ( $f$  es doblemente diferenciable sobre el dominio  $\mathbb{S}$ ) y acotada en  $\mathbb{S}$ . Para  $x \in \mathbb{S}$  tendremos que:

$$(\mathcal{L}f)(x) = \lim_{t \downarrow 0} \frac{(T_t f)(x) - f(x)}{t} = b(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x), \quad (5)$$

donde  $b(x)$  y  $\sigma^2(x)$  son los coeficientes de deriva y de difusión, respectivamente, definidos en 1.2.1 y 1.2.2 . (La demostración de este resultado puede encontrarse en [6]).

**Proposición 1.2.7:** Sea  $\{T_t : t > 0\}$  la familia de operadores de transición de un proceso de Markov con espacio de estados  $\mathbb{S}$ . Si  $f \in D_{\mathcal{L}}$ , entonces se satisface la siguiente ecuación diferencial parcial:

$$\frac{\partial}{\partial t}(T_t f)(x) = \mathcal{L}(T_t f), \quad t > 0, \quad x \in \mathbb{S}. \quad (6)$$

**Demostración:**

Demostremos que  $T_t f \in D_{\mathcal{L}}$ . Si  $f \in D_{\mathcal{L}}$ , ocupando el hecho de que  $T_t$  y  $T_s$  conmutan y que  $T_t$  es una contracción, tenemos que si  $s \downarrow 0$  entonces:

$$\left\| \frac{T_s(T_t f) - T_t f}{s} - T_t(\mathcal{L}f) \right\| = \left\| T_t \left( \frac{T_s f - f}{s} - \mathcal{L}f \right) \right\| \leq \left\| \frac{T_s f - f}{s} - \mathcal{L}f \right\| \rightarrow 0.$$

Por lo que  $T_t f \in D_{\mathcal{L}}$ , debido a que converge a  $T_t(\mathcal{L}f)$  y, por lo tanto, gracias a la definición de  $\mathcal{L}$ , tendremos (6).

Q.E.D.

**Nótese** que de la demostración de la proposición anterior tendremos que  $\mathcal{L}(T_t f) = T_t(\mathcal{L}f)$ , i.e.  $T_t$  y  $\mathcal{L}$  conmutan en  $D_{\mathcal{L}}$ .

Ahora, sea  $p(t; x, y)$  la densidad de transición de un proceso difusión  $\{X_t : t > 0\}$  y considérese  $f = \mathbb{1}_B$  en (4). Tendremos entonces que:

$$(T_{s+t}f)(x) = \int_B p(s+t; x, y) dy = \int_{\mathbb{S}} \left( \int_B p(t; z, y) dy \right) p(s; x, z) dz = \int_B \left( \int_{\mathbb{S}} p(t; z, y) p(s; x, z) dz \right) dy,$$

para cualquier conjunto de Borel  $B \subset \mathbb{S}$ . Lo cual implica que se satisface la *ecuación de Chapman-Kolmogorov*:

$$p(s+t; x, y) = \int_{\mathbb{S}} p(t; z, y) p(s; x, z) dz = (T_s f)(x),$$

donde  $f(z) = p(t; z, y)$ . Al aplicar (4) a esta  $f$ , obtendremos la *ecuación de retroceso de Kolmogorov* para la densidad de transición  $p(t; x, y)$ :

$$\frac{\partial p(t; x, y)}{\partial t} = b(x) \frac{\partial p(t; x, y)}{\partial x} + \frac{1}{2} \sigma^2(x) \frac{\partial^2 p(t; x, y)}{\partial x^2}. \quad (7)$$

Por otro lado, la *ecuación de evolución de Kolmogorov* nos otorga información acerca del futuro del proceso de difusión  $X_t$  cuando  $X_0$  posee una densidad de probabilidad inicial arbitraria. Supóngase que conocemos esta densidad inicial, le llamaremos  $g$ . Trataremos de obtener la ecuación de evolución de Kolmogorov en términos del operador adjunto de  $T_t$ . La densidad de  $X_t$  entonces estará dada por:

$$(T_t^* g)(y) = \int_{\mathbb{S}} g(x) p(t; x, y) dx.$$

El operador  $T_t^*$  transforma la densidad  $g$  en otra densidad de probabilidad. Este es el operador adjunto de  $T_t$  con respecto al producto interior  $\langle v, u \rangle = \int_{\mathbb{S}} v(x) u(x) dx$ . De hecho:

$$\langle T_t^* g, f \rangle = \int_{\mathbb{S}} (T_t^* g)(y) f(y) dy = \int_{\mathbb{S}} g(y) (T_t f)(y) dy = \langle g, T_t f \rangle.$$

Si  $f \in \mathcal{C}^2(\mathbb{S})$  y se anula afuera de un conjunto compacto de  $\mathbb{S}$ , podremos derivar con respecto a  $t$  en la ecuación anterior y, ocupando la conmutatividad de  $\mathcal{L}$  con  $T_t$  y (6), llegamos a que:

$$\left\langle \frac{\partial}{\partial t} T_t^* g, f \right\rangle = \left\langle g, \frac{\partial}{\partial t} T_t f \right\rangle = \langle g, T_t \mathcal{L} f \rangle = \langle T_t^* g, \mathcal{L} f \rangle = \int_{\mathbb{S}} (T_t^* g)(y) (\mathcal{L} f)(y) dy. \quad (8)$$

Si  $f, h \in \mathcal{C}^2(\mathbb{S})$  y se anulan afuera de un intervalo finito tendremos, al integrar por partes, que:

$$\begin{aligned} \langle h, \mathcal{L} f \rangle &= \int_{\mathbb{S}} h(y) \left[ b(y) f'(y) + \frac{1}{2} \sigma^2(y) f''(y) \right] dy \\ &= \int_{\mathbb{S}} \left[ -\frac{d}{dy} (b(y) h(y)) + \frac{d^2}{dy^2} (\sigma^2(y) h(y)) \right] f(y) dy = \langle \mathcal{L}^* h, f \rangle, \end{aligned}$$

donde  $\mathcal{L}^*$  es el *operador adjunto formal* del operador infinitesimal  $\mathcal{L}$ . Por lo tanto, aplicando lo anterior a (8), se tiene que:

$$\left\langle \frac{\partial}{\partial t} T_t^* g, f \right\rangle = \langle \mathcal{L}^* T_t^* g, f \rangle.$$

Dado que esta relación se satisface para una clase grande de funciones que se anulan afuera de un intervalo cerrado, tenemos la siguiente ecuación diferencial:

$$\frac{\partial}{\partial t} (T_t^* g)(y) = \mathcal{L}^* (T_t^* g)(y). \quad (9)$$

En otras palabras, la ecuación de evolución de Kolmogorov para la densidad de probabilidad de transición  $p(t; x, y)$  está dada por:

$$\frac{\partial p(t; x, y)}{\partial t} = -\frac{\partial}{\partial y} [b(y)p(t; x, y)] + \frac{1}{2} \frac{\partial^2}{\partial y^2} [\sigma^2(y)p(t; x, y)], \quad t > 0. \quad (10)$$

Nuevamente, la solución de esta ecuación dependerá del comportamiento del proceso de difusión en los puntos frontera del espacio de estados  $\mathbb{S}$ .

**Definición 1.2.8:** Sea  $\{X_t : t \geq 0\}$  un proceso de difusión con espacio de estados  $\mathbb{S} = [a, b]$  tal que  $-\infty \leq a < x < b \leq \infty$ . Denotaremos por  $\tau_y$  al primer tiempo de entrada o tiempo de paso al estado  $y$ , definido de la siguiente forma:

$$\tau_y = \inf\{t \geq 0 : X_t = y\}. \quad (11)$$

Para  $[c, d] \subset \mathbb{S}$  también denotaremos:

$$\tau_{c,d} = \inf\{t \geq 0 : X_t \notin (c, d)\} = \tau_c \wedge \tau_d. \quad (12)$$

De manera similar, considérese la filtración  $\{\mathcal{F}_t\}$ , decimos que  $\tau$  es un tiempo de paro de la filtración  $\{\mathcal{F}_t\}$  si  $\tau$  es un tiempo aleatorio de forma que para todo  $t \geq 0$  se satisface que  $\{\tau \leq t\} \in \mathcal{F}_t$ .

Existen algunos funcionales interesantes que podremos analizar al ocupar el operador infinitesimal  $\mathcal{L}$ . Estos funcionales jugarán un papel muy importante a la hora de analizar la *recurrencia* y *clasificación de fronteras*. Considérese el proceso de difusión  $\{X_t : t \geq 0\}$  donde  $X_0 = x$  y un intervalo cerrado  $[c, d] \subset \mathbb{S}$ ,  $c < d$ . Llamése

$$v(x) = v_{c,d}(x) = \mathbb{P}[\tau_c < \tau_d | X_0 = x], \quad c \leq x \leq d, \quad (13)$$

a la probabilidad de que el proceso alcance  $c$  antes que  $d$ , donde  $\tau_y$  al primer tiempo de entrada al estado  $y$  definido en (11). Para un tiempo pequeño  $h$ , al ocupar la ley de esperanza total y el teorema de Taylor (véase [26] y las referencias contenidas en los textos), asumiendo que  $v \in \mathcal{C}^2(\mathbb{S})$ , tendremos que:

$$v(x) = \mathbb{E}[v(X_h) | X_0 = x] + o(h) = v(x) + b(x)hv'(x) + \frac{1}{2}\sigma^2(x)hv''(x) + o(h).$$

Tenemos el siguiente resultado:

**Proposición 1.2.9:** Las probabilidades  $v(x)$  definidas en (13) son soluciones de la siguiente ecuación diferencial de segundo orden con condiciones iniciales:

$$\begin{aligned} \frac{1}{2}\sigma^2(x)v''(x) + b(x)v'(x) &= 0, \quad c < x < d, \\ v(c) &= 1, v(d) = 0. \end{aligned} \quad (14)$$

(La demostración de este resultado puede encontrarse en [26]).

Ahora vamos a estimar el *tiempo medio de salida*  $E(x)$  para alcanzar ya sea el estado  $c$  o  $d$  de un proceso de difusión que inicia en  $X_0 = x$ , donde  $x \in (c, d)$ , i.e.

$$E(x) = E_{c,d}(x) = \mathbb{E} [\tau_{c,d} | X_0 = x],$$

donde  $\tau_{c,d}$  se define en (12). Tal y como se hizo anteriormente, para un tiempo pequeño  $h$  y asumiendo que  $E \in \mathcal{C}^2(\mathbb{S})$ , tendremos que

$$E(x) = h + \mathbb{E} [E(X_h) | X_0 = x] + o(h) = h + E(x) + b(x)hE'(x) + \frac{1}{2}\sigma^2(x)hE''(x) + o(h).$$

Con lo cual tendremos el siguiente resultado:

**Proposición 1.2.10:** *El tiempo medio de salida  $E(x)$  es una solución de la siguiente ecuación diferencial de segundo orden con condiciones iniciales*

$$\begin{aligned} \frac{1}{2}\sigma^2(x)E''(x) + b(x)E'(x) &= -1, \quad c < x < d, \\ E(c) = E(d) &= 0. \end{aligned} \tag{15}$$

(La demostración de este resultado puede encontrarse en [26]).

Denótese

$$\rho_{xy} = \mathbb{P} [\tau_y < \infty | X_0 = x]. \tag{16}$$

**Definición 1.2.11:** *Ocupando la notación anterior, tendremos que:*

- *Un estado  $y$  es recurrente si  $\rho_{xy} = 1$  para todo  $x \in \mathbb{S}$ , de forma que  $\rho_{yx} > 0$ , y transitivo de otra forma.*
- *Un proceso de difusión recurrente es recurrente positivo si  $\mathbb{E}[\tau_y | X_0 = x] < \infty$  para todo  $x, y \in \mathbb{S}$  y nulo recurrente de otra forma.*

Si todos los estados de  $\mathbb{S}$  son recurrentes, entonces decimos que *el proceso de difusión es recurrente*. En casos donde el proceso de difusión es recurrente positivo, existirá una *distribución estacionaria* o *distribución invariante*  $\pi(y)$  que satisface:

$$\pi(y)dy = (T_t^* \pi)(y)dy = \int_{\mathbb{S}} \pi(x)p(t; x, dy)dx := \mathbb{P}_\pi[X_t \in dy], \quad t > 0. \tag{17}$$

Para cadenas de Markov, esta distribución invariante satisface que:

$$\lim_{t \rightarrow \infty} p(t; x, y) = \pi(y).$$

Por lo tanto, al ocupar lo anterior en la ecuación de evolución de Kolmogorov (10) o (9), obtendremos que la distribución invariante satisface que  $\mathcal{L}^* \pi(y) = 0$ , i.e.

$$\frac{1}{2} \frac{d^2}{dy^2} [\sigma^2(y)\pi(y)] - \frac{d}{dy} [b(y)\pi(y)] = 0, \quad y \in \mathbb{S},$$

siempre y cuando todos los puntos de la frontera no puedan ser alcanzados desde cualquier punto dentro del espacio de estados. En otras palabras:

$$\frac{d}{dy} [\sigma^2(y)\pi(y)] = 2b(y)\pi(y), \quad y \in \mathbb{S}.$$

En el estudio de los procesos de difusión usualmente, al trabajar con un proceso  $X(t)$  con  $X(0) = x$ , se analiza la función  $(T_t f)(x) := \mathbb{E}[f(X(t)) | X(0) = x]$ , (definida a través de  $f \in \mathbb{B}(\mathbb{S})$ ), la cual es de hecho el operador de transición  $T_t$  (véase **Definición 1.2.4**).

**Definición 1.2.12:** Un proceso  $X(t)$  es de Feller si la función  $(T_t f)(\cdot)$  es continua para todo  $t \geq 0$  y  $\lim_{t \downarrow 0} (T_t f)(x) = f(x)$  para cualquier función  $f$  medible, continua y acotada.

**Definición 1.2.13** Considérese la función  $(T_t f)(x) := \mathbb{E}[f(X(t)) | X(0) = x]$  previamente definida y también considérese la **Definición 1.2.12**. Diremos que el proceso  $X(t)$  satisface la propiedad fuerte de Feller si satisface la propiedad de Feller y si la función  $(T_t f)(\cdot)$  es continua para todo  $t > 0$  y para cualquier función  $f$  acotada y medible.

Todos estos conceptos y resultados obtenidos en esta sección serán redefinidos y adaptados a lo largo de los Capítulos II y III para ajustarlos a los procesos de difusión cambiante, por lo que se reutilizará la misma notación.

### 1.3 Polinomios ortogonales y medidas matriciales

En esta sección ocuparemos como referencia principal [11].

Comenzamos definiendo el concepto de polinomios ortogonales.

**Definición 1.3.1:** Sea  $\psi$  una medida de Borel real positiva, definimos los momentos de  $\psi$  como:

$$\mu_n := \int_{\mathbb{R}} x^n d\psi(x), \quad n \in \mathbb{N}_0. \quad (18)$$

Supondremos siempre que existen y que son finitos. Normalizaremos la medida de tal forma que  $\mu_0 = 1$  con el objeto de que  $\mu$  sea una medida de probabilidad.

**Definición 1.3.2:** El espacio de Hilbert  $L^2_{\psi}(\mathbb{R})$  con producto interno:

$$(f, g)_{\psi} = \int_{\mathbb{R}} f(x)g(x)d\psi(x),$$

está definido para toda función real medible  $f$  tal que  $\|f\|_{\psi}^2 := (f, f)_{\psi} < \infty$ , es decir  $f$  tiene norma finita.

En el caso de que el soporte de esta medida sea  $S \subset \mathbb{R}$ , denotaremos este espacio como  $L^2_{\psi}(S)$ . Si  $S$  fuera un conjunto numerable, por ejemplo  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ , el espacio se denotará por  $\ell^2_{\psi}(S)$ .

**Definición 1.3.3:** Decimos que  $(p_n(x))_{n \in \mathbb{N}}$  es una sucesión de polinomios si cada elemento es un polinomio de grado  $n$  de variable real  $x$ . Una sucesión de polinomios es mónica si el coeficiente de mayor grado de cada polinomio es 1.

**Definición 1.3.4:** Una sucesión de polinomios  $(p_n)_n$  es ortogonal con respecto a una medida de Borel  $\psi$  si

$$(p_n, p_m)_{\psi} = \int_{\mathbb{R}} p_n(x)p_m(x)d\psi(x) = d_n^2 \delta_{nm},$$

tal que  $d_n^2 := (p_n, p_n)_{\psi} = \|p_n\|_{\psi}^2 > 0$ .

En caso de que la norma sea idéntica a 1, decimos que la sucesión de polinomios es ortonormal y la denotamos por  $(P_n)_n$ . En caso de tener una sucesión de polinomios ortogonales mónicos, utilizaremos la notación  $(\hat{P}_n)_n$  y las normas se denotarán por  $\zeta_n = \|\hat{P}_n\|_{\psi}^2$ .

Dada una medida de Borel  $\psi$  en  $\mathbb{R}$  con soporte infinito y momentos finitos, siempre será posible construir una sucesión de polinomios ortogonales por el conocido procedimiento de Gram-Schmidt.

Otro método para generar polinomios ortogonales ocupa **recurrencia**: supóngase la existencia de una sucesión de polinomios ortogonales  $(p_n)_n$ . El polinomio  $xp_n(x)$  es de grado  $n+1$  y puede ser expresado como una combinación lineal de los primeros  $n+1$  polinomios:

$$xp_n(x) = \sum_{j=0}^{n+1} d_{n,j} p_j(x).$$

Multiplicando por  $p_k(x)$ , evaluando el producto interior y ocupando las relaciones de ortogonalidad tendremos que  $d_{n,j} = 0 : j = 0, 1, \dots, n-2$ . Por lo que sólo los últimos tres términos sobreviven y por ende *toda familia de polinomios ortogonales satisface una relación de recurrencia a tres términos*:

$$xp_n(x) = a_n p_{n+1}(x) + b_n p_n(x) + c_n p_{n-1}(x), \quad n \geq 0, \quad p_{-1} = 0, \quad (19)$$

donde:

$$a_n = \frac{(xp_n, p_{n+1})_\psi}{(p_{n+1}, p_{n+1})_\psi}, \quad b_n = \frac{(xp_n, p_n)_\psi}{(p_n, p_n)_\psi}, \quad c_n = \frac{(xp_n, p_{n-1})_\psi}{(p_{n-1}, p_{n-1})_\psi}.$$

Nótese que  $b_n \in \mathbb{R}$ . Además, tenemos para la familia de polinomios ortonormales  $P_n(x)$  que  $a_n = \sqrt{\frac{\zeta_{n+1}}{\zeta_n}} > 0$  y que  $c_n = (xP_n, P_{n-1})_\psi = (P_n, xP_{n-1})_\psi = a_{n-1}$ . Por lo tanto la sucesión de polinomios ortonormales  $(P_n)_n$ , satisface una relación de recurrencia a tres términos de la forma:

$$xP_n(x) = a_n P_{n+1}(x) + b_n P_n(x) + a_{n-1} P_{n-1}(x), \quad a_n > 0, \quad b_n \in \mathbb{R}, \quad (20)$$

mientras que la familia mónica  $(\hat{P}_n)_n$  satisface que:

$$x\hat{P}_n(x) = \hat{P}_{n+1}(x) + \alpha_n \hat{P}_n(x) + \beta_n \hat{P}_{n-1}(x), \quad \hat{P}_0(x) = 1, \quad \hat{P}_1(x) = x - \alpha_0, \quad (21)$$

de forma tal que  $\alpha_{n-1} \in \mathbb{R}, \beta_n > 0 : n \geq 1$ . Estos coeficientes y los coeficientes de la familia ortonormal satisfacen:

$$a_n = \sqrt{\frac{\zeta_{n+1}}{\zeta_n}}, \quad \alpha_n = b_n, \quad \beta_n = \frac{\zeta_n}{\zeta_{n-1}}.$$

Nótese también que  $\zeta_n = \beta_n \cdots \beta_1$ .

Otra forma de escribir esta relación de recurrencia es a través de su forma matricial. Sea  $P(x) = (P_0(x), P_1(x), \dots)^T$  el vector columna de los polinomios ortogonales, tenemos que  $xP(x) = JP(x)$ , donde  $J$  es la matriz tridiagonal simétrica:

$$J = \begin{pmatrix} b_0 & a_0 & 0 & 0 & 0 & \cdots \\ a_0 & b_1 & a_1 & 0 & 0 & \cdots \\ 0 & a_1 & b_2 & a_2 & 0 & \cdots \\ 0 & 0 & a_2 & b_3 & a_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}. \quad (22)$$

A esta matriz se le llama *matriz de Jacobi*. El *Teorema de Favard* o *Teorema Espectral para Polinomios Ortogonales*, garantiza que dada una familia de polinomios que satisfaga y esté definida por (20), existe una medida de Borel positiva bajo la cual  $(P_n)_n$  son ortonormales.

Mostraremos un ejemplo de polinomios ortogonales con respecto a una medida, en este caso hablamos de los *polinomios de Jacobi*. Este ejemplo, denominado como una de las familias clásicas, se usará en el Capítulo III en relación con el método de Wright-Fisher.

### Ejemplo 1.3.5: Polinomios de Jacobi

La forma más general de definir los Polinomios de Jacobi  $\{P_n^{(\alpha,\beta)}\}_{n \in \mathbb{N}}$  es a través de la denominada fórmula de Rodrigues:

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{n!2^n} (1-x)^{-\alpha} (1+x)^{-\beta} \frac{d^n}{dx^n} ((1-x)^{\alpha+n} (1+x)^{\beta+n}), \quad n \geq 0.$$

Estos polinomios son ortogonales con respecto a la medida  $\mu(x) = (1-x)^\alpha (1+x)^\beta$ ,  $\alpha, \beta > -1$ , con soporte en  $x \in [-1, 1]$  y satisfacen la siguiente ecuación diferencial de segundo orden:

$$(1-x^2) \left[ P_n^{(\alpha,\beta)} \right]''(x) + [\beta - \alpha - (\beta + \alpha + 2)x] \left[ P_n^{(\alpha,\beta)} \right]'(x) + n(n + \alpha + \beta + 1) P_n^{(\alpha,\beta)}(x) = 0, \quad n \geq 0.$$

Los polinomios de Jacobi satisfacen la siguiente relación de recurrencia a tres términos para  $n \geq 0$  con  $P_{-1}^{(\alpha,\beta)} := 0, P_0^{(\alpha,\beta)} := 1$ :

$$\begin{aligned} x P_n^{(\alpha,\beta)}(x) &= \frac{2(n+1)(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} P_{n+1}^{(\alpha,\beta)}(x) \\ &+ \frac{\beta^2 - \alpha^2}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)} P_n^{(\alpha,\beta)}(x) + \frac{2(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)} P_{n-1}^{(\alpha,\beta)}(x). \end{aligned}$$

Por lo que la matriz de Jacobi para este caso es

$$J = \begin{pmatrix} \frac{\beta-\alpha}{\beta+\alpha+2} & \frac{2}{\beta^2-\alpha^2} & 0 & 0 & \dots \\ \frac{2(1+\alpha)(1+\beta)}{(2+\alpha+\beta)(2+\alpha+\beta+1)} & \frac{\beta^2-\alpha^2}{(2+\alpha+\beta)(2+\alpha+\beta+2)} & \frac{4(1+\alpha+\beta+1)}{(2+\alpha+\beta+1)(2+\alpha+\beta+2)} & 0 & \dots \\ 0 & \frac{2(2+\alpha)(2+\beta)}{(4+\alpha+\beta)(4+\alpha+\beta+1)} & \frac{\beta^2-\alpha^2}{(4+\alpha+\beta)(4+\alpha+\beta+2)} & \frac{6(2+\alpha+\beta+1)}{(4+\alpha+\beta+1)(4+\alpha+\beta+2)} & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}.$$

Otra manera de definir los polinomios de Jacobi es considerándolos en el intervalo  $[0, 1]$ . En este caso los denotaremos por  $\{Q_n^{(\alpha,\beta)}\}_{n \in \mathbb{N}}$ , realizando el cambio de variable  $y = \frac{1+x}{2}$  de forma que el punto  $x = -1$  sea  $y = 0$  y el punto  $x = 1$  sea  $y = 1$ , lo cual también cambia el parámetro  $\alpha$  por  $\beta$ . En este caso, todas las propiedades que satisfacen los polinomios de Jacobi sobre  $[-1, 1]$  pueden extenderse de manera natural. En particular, si  $n \geq 0$ , los polinomios de Jacobi sobre  $[0, 1]$  satisfacen la ecuación diferencial de segundo orden

$$y(1-y) \left[ Q_n^{(\alpha,\beta)} \right]''(y) + [(1-y)(\alpha+1) - y(\beta+1)] \left[ Q_n^{(\alpha,\beta)} \right]'(y) + n(n + \alpha + \beta + 1) Q_n^{(\alpha,\beta)}(y) = 0.$$

Además, los polinomios satisfacen la siguiente relación de recurrencia a tres términos para  $n \geq 0$  con  $Q_n^{(\alpha,\beta)}(1) = 1$ ,

$$y Q_n^{(\alpha,\beta)}(y) = a_n Q_{n+1}^{(\alpha,\beta)}(y) + b_n Q_n^{(\alpha,\beta)}(y) + c_n Q_{n-1}^{(\alpha,\beta)}(y), \quad n \geq 0,$$

donde los coeficientes  $a_n, b_n, c_n$  pueden escribirse de la siguiente forma:

$$\begin{aligned}
a_n &= \frac{(n + \beta + 1)(n + 1 + \alpha + \beta)}{(2n + \alpha + \beta + 1)(2n + 2 + \alpha + \beta)}, \quad n \geq 0, \\
b_n &= \frac{(n + \beta + 1)(n + 1)}{(2n + \alpha + \beta + 1)(2n + 2 + \alpha + \beta)} + \frac{(n + \alpha)(n + \alpha + \beta)}{(2n + \alpha + \beta + 1)(2n + \alpha + \beta)}, \quad n \geq 0, \\
c_n &= \frac{n(n + \alpha)}{(2n + \alpha + \beta + 1)(2n + \alpha + \beta)}, \quad n \geq 1.
\end{aligned}$$

Obsérvese que la ecuación diferencial anterior corresponde al modelo de difusión de Wright-Fisher que involucra solo efectos de mutación (véase [26]), mientras que la matriz de Jacobi generada por los coeficientes  $a_n, b_n, c_n$  es una matriz estocástica (i.e una cadena de nacimiento y muerte a tiempo discreto). Para más detalle véase [11].

Para finalizar esta sección, extenderemos ahora el concepto de ortogonalidad a funciones de valores matriciales. Sea  $\mathcal{M}_N$  el anillo de todas las matrices complejas de dimensión  $N \times N$  (usualmente denotado por  $\mathbb{C}^{N \times N}$  o  $\mathbb{R}^{N \times N}$  para el caso real), vamos a denotar por  $a^*$  a la matriz conjugada transpuesta (para el caso real ocuparemos  $a^t$  para denotar a la matriz transpuesta). De igual forma denotamos  $\|a\| = \sup_{z \in \mathbb{C}^N} \frac{\|az\|_2}{\|z\|_2}$  como la norma euclidiana de  $a \in \mathcal{M}_N$  visto como un operador lineal de  $\mathbb{C}^N$  en  $\mathbb{C}^N$ . Considérese  $\mathcal{P}$  como el conjunto de polinomios de variable  $z \in \mathbb{C}$  con coeficientes en  $\mathcal{M}_N$ . Sea  $n = 0, 1, \dots$ , entonces  $\mathcal{P}_n$  denota al conjunto de polinomios de grado menor o igual a  $n$ .

**Definición 1.3.6:** Una medida matricial  $\mu$  sobre  $\mathbb{R}$  (o si se prefiere  $\mathbb{C}$ ) es la asignación de una matriz  $N \times N$  semi-definida positiva  $\mu(X)$  a cualquier conjunto de Borel  $X$ . Es decir:

$$\forall X \in \mathbb{B}, \quad \mu(X) \in \mathcal{M}_N : \forall x \in \mathbb{R}^N \quad x^* \mu(X) x \geq 0. \quad (23)$$

Normalmente, normalizaremos a la medida  $\mu$  de forma tal que  $\mu(\mathbb{R}) = \text{Id}_{N \times N} = 1$  (o  $\mu(\mathbb{C}) = \text{Id}_{N \times N} = 1$ ).

Dada una medida matricial  $\mu$  cualquiera, le asignaremos una medida escalar  $\mu_{tr}(X) = \text{Tr}(\mu(X))$ , donde  $\text{Tr}(\text{Id}_{N \times N}) = \text{Tr}(1) = N$ . Nuevamente, normalizamos  $\mu_{tr}(\mathbb{R}) = N$ . Cuando  $N = 1$ , tendremos una medida de Borel real y positiva.

Ahora, si aplicamos el *Teorema de Radon-Nikodym* a cada entrada o elemento de la medida  $\mu$ , podemos garantizar la existencia de una función matricial semi-definida positiva  $M(x) = [M_{ij}(x)]$  tal que:

$$d\mu_{ij}(x) = M_{ij}(x) d\mu_{tr}(x),$$

donde  $\text{Tr}(M(x)) = 1$ . Por el contrario, cualquier medida escalar  $\mu_{tr}$  tal que  $\mu_{tr}(\mathbb{R}) = N$ , con función matricial semi-definida positiva  $M(x)$  que cumpla con  $\text{Tr}(M(x)) = 1$  define a una medida matricial  $\mu$  normalizada tal que  $\mu(\mathbb{R}) = \text{Id}_{N \times N} = 1$ .

**Definición 1.3.7:** Sea  $\mu$  una medida matricial, el producto interno  $\langle \cdot, \cdot \rangle_\mu$  con respecto a  $\mu$  para cualesquiera  $P, Q \in \mathcal{P}$ , está definido como:

$$\langle P, Q \rangle_\mu = \int_{\mathbb{R}} Q^*(x) d\mu(x) P(x) \in \mathbb{C}^{N \times N}. \quad (24)$$

Para cualesquiera  $A, B \in \mathcal{M}_N$  y  $P, Q, R \in \mathcal{P}$  se tienen las siguientes propiedades:

- Linealidad:

$$\begin{aligned}
\langle PA + QB, R \rangle_\mu &= \int_{\mathbb{R}} R^*(x) d\mu(x) (PA + QB)(x) = \int_{\mathbb{R}} R^*(x) d\mu(x) P(x) A + \int_{\mathbb{R}} R^*(x) d\mu(x) Q(x) B \\
&= \langle P, R \rangle_\mu A + \langle Q, R \rangle_\mu B.
\end{aligned}$$



- Conjugación:

$$\langle Q, P \rangle_\mu^* = \left( \int_{\mathbb{R}} P^*(x) d\mu(x) Q(x) \right)^* = \int_{\mathbb{R}} Q^*(x) d\mu(x) P(x) = \langle P, Q \rangle_\mu.$$

- Positividad:

$$\langle P, P \rangle_\mu \geq 0, \text{ además } \langle P, P \rangle_\mu = 0 \iff P = 0.$$

**Definición 1.3.8:** Sea  $F$  una función matricial. A la función  $F$  le asignaremos el escalar  $\|F\|_\mu$  definido por:

$$\|F\|_\mu = \sqrt{\text{Tr}(\langle F, F \rangle_\mu)}. \quad (25)$$

Decimos que  $\|\cdot\|_\mu$  es una semi-norma debido a que  $\langle\langle F, G \rangle\rangle_\mu = \text{Tr}(\langle F, G \rangle_\mu)$  es una forma cuasi-lineal (en el aspecto de que esta forma no satisface la definición de linealidad en su totalidad) que a su vez es semi-definida positiva y  $\|\cdot\|_\mu$  es la ‘norma’ asociada a este ‘producto interno’. De hecho, el producto interno  $\langle\langle F, G \rangle\rangle_\mu$ , satisface la desigualdad de Cauchy-Schwarz:

$$|\text{Tr}(\langle F, G \rangle_\mu)| \leq \|F\|_\mu \|G\|_\mu.$$

**Definición 1.3.9:** El conjunto de clases equivalentes determinadas por  $F \sim G \iff \|F - G\|_\mu = 0$  define un espacio de Hilbert  $\mathcal{H}$ , cuyo producto interno es  $\langle\langle F, G \rangle\rangle_\mu = \text{Tr}(\langle F, G \rangle_\mu)$ .

**Definición 1.3.10:** Sea  $F(x)$  una función matricial, definimos la norma del determinante de  $F$  como:

$$|F| = \sqrt{\det(\langle F, F \rangle_\mu)},$$

la cual es ‘más fuerte’ que  $\|F\|_\mu$  en el sentido de que  $|F| > 0 \implies \|F\|_\mu > 0$ , pero el contrario no necesariamente es cierto.

**Definición 1.3.11:** Un conjunto  $\{\psi_j\}_{j=1}^n$  en  $\mathcal{H}$  (véase **Definición 1.3.9**) es ortonormal si y solo si:

$$\langle\langle \psi_j, \psi_k \rangle\rangle_\mu = \delta_{jk} \mathbf{I}_{N \times N}.$$

Supóngase que  $n < \infty$ , sean  $a_1, \dots, a_n, b_1, \dots, b_n \in \mathcal{M}_N$ , por lo que  $\sum_{j=1}^n \psi_j a_j$  y  $\sum_{j=1}^n \psi_j b_j$  están correctamente definidos y (ocupando las propiedades del producto por matriz constante de  $\langle\langle \cdot, \cdot \rangle\rangle_\mu$ ) tendremos que:

$$\left\langle \sum_{j=1}^n \psi_j a_j, \sum_{j=1}^n \psi_j b_j \right\rangle_\mu = \sum_{j=1}^n b_j^* a_j.$$

Denotamos al conjunto de posibles combinaciones lineales de  $\{\psi_j\}_{j=1}^n$  como  $\mathcal{H}_{(\psi_j)} = \{\sum_{j=1}^n \psi_j a_j : \{a_j\}_{j=1}^n \subset \mathcal{M}_N, n \in \mathbb{N}\}$ , el cual es un subespacio vectorial de  $\mathcal{H}$  (sobre  $\mathbb{C}$ ) de dimensión  $nN^2$ .

Ahora, sea  $F \in \mathcal{H}$ , la proyección ortogonal de  $F$  con respecto al producto interno  $\langle\langle \cdot, \cdot \rangle\rangle_\mu$ , es de la forma:

$$\pi_{(\psi_j)}(F) = \sum_{j=1}^n \psi_j \langle\langle \psi_j, F \rangle\rangle_\mu.$$

Nótese que por la definición de  $\pi_{(\psi_j)}(F)$  y por la ortonormalidad, podemos obtener con facilidad el *Teorema de Pitágoras en  $\mathcal{H}$* :

$$\langle F, F \rangle_\mu = \langle F - \pi_{(\psi_j)}(F), F - \pi_{(\psi_j)}(F) \rangle_\mu + \sum_{j=1}^n \langle \psi_j, F \rangle_\mu^* \langle \psi_j, F \rangle_\mu.$$

Si hacemos tender  $n$  al infinito:  $n \rightarrow +\infty$ , tendremos que:

$$\lim_{n \rightarrow +\infty} \left[ \sum_{j=1}^n \langle \psi_j, F \rangle_\mu^* \langle \psi_j, F \rangle_\mu \right] \leq \langle F, F \rangle_\mu,$$

y también tendremos la convergencia de la proyección ortogonal:

$$\lim_{n \rightarrow +\infty} \left[ \sum_{j=1}^n \psi_j \langle \psi_j, F \rangle_\mu \right] = \pi_{(\psi_j)}(F).$$

**Definición 1.3.12:** *El conjunto de funciones ortonormales  $\{\psi_j\}_{j=1}^n$  es completo en  $\mathcal{H}$  cuando  $\mathcal{H}_{(\psi_j)} = \mathcal{H}$ . Nótese que en este caso, se satisface que:*

$$\lim_{n \rightarrow +\infty} \left[ \sum_{j=1}^n \langle \psi_j, F \rangle_\mu^* \langle \psi_j, F \rangle_\mu \right] = \langle F, F \rangle_\mu,$$

y:

$$\pi_{(\psi_j)}(F) = \lim_{n \rightarrow +\infty} \left[ \sum_{j=1}^n \psi_j \langle \psi_j, F \rangle_\mu \right] = F.$$

Por ende, para bases ortonormales, tendremos la **igualdad de Parseval**:

$$\sum_{j=1}^{\infty} \langle \psi_j, F \rangle_\mu^* \langle \psi_j, F \rangle_\mu = \langle F, F \rangle_\mu.$$

Además:

$$\|F\|_\mu^2 = \sum_{j=1}^{\infty} \text{Tr}(\langle \psi_j, F \rangle_\mu^* \langle \psi_j, F \rangle_\mu).$$

## 1.4 Resultados adicionales

En esta última sección del Capítulo I se enlistarán una serie de definiciones y resultados que serán de gran utilidad para demostrar y/o analizar los conceptos que se tratarán en los Capítulos II y III.

**Definición 1.4.1:** *Un proceso estocástico  $\{X_t\}_{t \geq 0}$  es una cadena de Markov a tiempo continuo que toma valores en un espacio de estados  $S = \{1, \dots, N\}$ , (usualmente finito), si existe una matriz  $Q = [q_{ij}]_{i,j \in S}$ , denominada operador infinitesimal del proceso, la cual satisface que:*

$$q_{ij} = \begin{cases} \lim_{t \downarrow 0} \frac{P_{ij}(t)}{t} = \lim_{t \downarrow 0} \frac{\mathbb{P}[X_t = i | X_0 = j]}{t}, & \text{cuando } i \neq j, \\ \lim_{t \downarrow 0} \frac{P_{ii}(t) - 1}{t} = \lim_{t \downarrow 0} \frac{\mathbb{P}[X_t = i | X_0 = i] - 1}{t}, & \text{en otro caso.} \end{cases}$$

Además, se tiene que  $q_{ii} \leq 0$ ,  $\forall_{i \neq j} q_{ij} \geq 0$  y  $-q_{ii} = \sum_{i \neq j} q_{ij}$ .

Empezando en el estado  $i$ , el proceso se mantiene ahí un lapso aleatorio distribuido como una variable aleatoria exponencial de media  $-\frac{1}{q_{ii}}$ . Seguido se mueve a otro estado seleccionado aleatoriamente  $j$  con probabilidad  $-\frac{q_{ij}}{q_{ii}}$ . Este procedimiento se repite para el estado  $j$  y para todos los estados subsecuentes.

A lo largo de este texto asumiremos que el proceso  $\{X_t\}_{t \geq 0}$  es homogéneo en el tiempo, es decir, para cualesquiera  $t, s > 0$ ,  $t - s > 0$ ,  $i, j \in S$  tendremos que  $\mathbb{P}[X_t = i | X_s = j] = \mathbb{P}[X_{t-s} = i | X_0 = j]$ . En este caso siempre será posible definir la función  $P_{ij}(t) = \mathbb{P}[X_t = i | X_0 = j]$  a la que denominaremos función de probabilidad de transición. Estas probabilidades pueden ser escritas en la denominada matriz de probabilidad de transición:

$$P(t) = \begin{bmatrix} P_{00}(t) & P_{01}(t) & P_{02}(t) & \cdots \\ P_{10}(t) & P_{11}(t) & P_{12}(t) & \cdots \\ P_{20}(t) & P_{21}(t) & P_{22}(t) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

La matriz de probabilidad de transición  $P(t)$  satisface las siguientes propiedades:

- 1)  $P_{ij}(t) \geq 0$  para cualesquiera  $i, j \in S$  y  $P_{ij}(0) = \delta_{ij}$ , la delta de Kronecker.
- 2)  $\sum_{j \in S} P_{ij}(t) \leq 1$  para cualesquiera  $t \geq 0$ ,  $i \in S$ . Diremos que el proceso  $P_{ij}(t)$  es honesto si  $\sum_{j \in S} P_{ij}(t) = 1$  para cualesquiera  $t \geq 0$ ,  $i \in S$ , y diremos que es deshonesto en otro caso.
- 3) Para  $i, j \in S$  y para cualesquiera  $s, t \geq 0$ , tendremos las ecuaciones de Chapman-Kolmogorov (o bien, la propiedad de semigrupo):

$$P_{ij}(s+t) = \sum_{k \in S} P_{ik}(s)P_{kj}(t).$$

Por ejemplo, para el caso donde las transiciones entre estados ocurran únicamente entre estados adyacentes, es decir  $i \rightarrow i+1$  o  $i \rightarrow i-1$ , entonces diremos que  $\{X_t\}_{t \geq 0}$  es una *proceso de nacimiento y muerte a tiempo continuo*. (Para más información consultar [1]).

**Definición 1.4.2:** Un proceso estocástico  $X(t)$  con trayectorias continuas por la derecha es continuo en probabilidad en  $t$  si para algún  $\eta > 0$ ,

$$\lim_{\Delta \rightarrow 0} \mathbb{P}[|X(t+\Delta) - X(t)| \geq \eta] = 0.$$

Es continuo en media cuadrática en  $t$  si

$$\lim_{\Delta \rightarrow 0} \mathbb{E}[|X(t+\Delta) - X(t)|^2] = 0.$$

**Definición 1.4.3:** Sea  $f : \mathbb{R}^r \mapsto \mathbb{R}$ ,  $r \in \mathbb{N}$  una función, decimos que  $f$  es:

- a) Hölder continua de exponente  $\alpha \geq 1$  si

$$\sup_{x \neq y \in \mathbb{R}^r} \frac{|f(x) - f(y)|}{\|x - y\|_\infty^\alpha} < \infty,$$

donde  $\|\cdot\|_\infty$  es la norma uniforme o la norma del supremo de  $\mathbb{R}^r$ .

b) Localmente Hölder continua de exponente  $\alpha \geq 1$  si

$$\sup_{x \neq y \in K} \frac{|f(x) - f(y)|}{\|x - y\|_\infty^\alpha} < \infty,$$

para todo subconjunto compacto  $K \subset \mathbb{R}^r$ .

c) Hölder continua en  $x_0 \in \mathbb{R}^r$  de exponente  $\alpha \geq 1$  si

$$\sup_{x \in U} \frac{|f(x) - f(x_0)|}{\|x - x_0\|_\infty^\alpha} < \infty,$$

para alguna vecindad  $U \subset \mathbb{R}^r$  de  $x_0$  tal que  $x_0 \notin U$ .

**Proposición 1.4.4 [Desigualdad de Hölder]:** Sean  $p, q > 1$  de forma que  $\frac{1}{p} + \frac{1}{q} = 1$ , entonces para cualesquiera dos funciones  $f, g : S = (a, b) \mapsto \mathbb{R}$  real-medibles se satisface que

$$\int_a^b |f(x)g(x)|dx \leq \left[ \int_a^b |f(x)|^p dx \right]^{\frac{1}{p}} \left[ \int_a^b |g(x)|^q dx \right]^{\frac{1}{q}}.$$

(La demostración de este resultado puede encontrarse en [30]).

**Proposición 1.4.5 [Desigualdad de Chebyshev]:** Sea  $X$  una variable aleatoria cualquiera con media y varianza finita,  $\mu$  y  $\sigma^2$  respectivamente. Entonces para cualquier  $a \in \mathbb{R}$ ,

$$\mathbb{P}[|X - \mu| \geq a] \leq \frac{\sigma^2}{a^2}.$$

(La demostración de este resultado puede encontrarse en [31]).

**Proposición 1.4.6 [Desigualdad de Gronwall]:** Sea  $I$  un intervalo real de la forma  $[a, \infty)$  o  $[a, b]$  o  $[a, b)$  tal que  $a < b$ . Sean  $f$  y  $g$  funciones continuas real-valuadas definidas sobre  $I$ . Si  $g$  es diferenciable en el interior  $I^\circ$  de  $I$  y satisface la desigualdad diferencial

$$g'(t) \leq f(t)g(t), \quad t \in I^\circ,$$

entonces  $g$  está acotada por la solución de la siguiente ecuación diferencial  $v'(t) = f(t)v(t)$ , es decir:

$$g(t) \leq f(a)e^{\int_a^t f(s)ds}, \quad \text{para todo } t \in I.$$

(La demostración de este resultado puede encontrarse en [21]).

**Proposición 1.4.7 [Desigualdad de Harnack]:** Sea  $\Omega$  un dominio acotado de  $\mathbb{R}^r$ . Sea  $\Gamma \subset \Omega$  un conjunto cerrado. Existe una constante  $K_1 > 0$  que depende únicamente de  $r, n$ , el diámetro de  $\Omega$ , la distancia de  $\Gamma$  a  $\partial\Omega$  y una constante  $\lambda$ , tal que para cualquier función  $\mathcal{L}$ -armónica  $u$  en  $\Omega$ , entonces

$$\sup_{x \in \Gamma} \{u_i(x)\} \leq K_1 \max_{1 \leq k \leq n} \inf_{x \in \Gamma} \{u_k(x)\}, \quad \text{para todo } i \in \{1, \dots, n\}.$$

**Nota:** decimos que  $f \geq 0$  una función  $\mathcal{L}$ -armónica en  $D$  si, dado un proceso de difusión cambiante  $(X(\cdot), Y(\cdot))$  tal que  $X(t) \in D \subset \mathbb{R}$  e  $Y(t) \in \mathbb{S} \subset \mathbb{N}$  (el cual se definirá en el Capítulo II) existe un tiempo de paro  $\tau$  (véase **Definición 1.2.8**) tal que para todo  $x \in D$  e  $i \in \mathbb{S}$  se satisface que

$$f(x, i) = \mathbb{E}[f(X(\tau), Y(\tau)) | X(0) = x, Y(0) = i]$$

(La demostración y definiciones de este resultado puede encontrarse en [3]).

**Definición 1.4.8 [Propiedad de Markov]:** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad con una filtración  $\{\mathcal{F}_s : s \in I\}$  totalmente ordenada con respecto a algún índice  $I$  (en este caso  $I = \mathbb{R}$ ),  $X = \{X_t | t \geq 0\}$  un proceso estocástico a tiempo continuo con espacio de estados  $\mathbb{S}$ , decimos que  $X$  satisface la propiedad de Markov si para cualesquiera  $A \subset \mathbb{S}$ ,  $s, t \in I$  tal que  $s < t$ ,

$$\mathbb{P}[X_t \in A | \mathcal{F}_s] = \mathbb{P}[X_t \in A | X_s].$$

**Definición 1.4.9 [Propiedad fuerte de Markov]:** Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad con una filtración  $\{\mathcal{F}_t\}_{t \geq 0}$ . Para cualquier tiempo de paro  $\tau \in \Omega$  (**Definición 1.2.8**), podemos definir

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : \forall t \geq 0, \{\tau \leq t\} \cap A \in \mathcal{F}_t\}.$$

Decimos que  $X = \{X_t | t \geq 0\}$ , un proceso estocástico a tiempo continuo con espacio de estados  $\mathbb{S}$ , satisface la propiedad fuerte de Markov si para cualquier tiempo de paro  $\tau$ , condicionado a que  $\tau < \infty$ , tenemos que para todo  $t \geq 0$ ,  $X_{\tau+t}$  es independiente de  $\mathcal{F}_\tau$  dado  $X_\tau$ .

**Definición 1.4.10 [Condición de crecimiento lineal]:** Sea  $f : \mathbb{R} \mapsto \mathbb{R}$  una función, decimos que  $f$  satisface la condición de crecimiento lineal si para todo  $x \in \mathbb{R}$  existen dos constantes  $a, b > 0$  tales que:

$$|f(x)| \leq a|x| + b.$$

Por otro lado, decimos que  $f$  satisface la condición de crecimiento lineal local si existe un subconjunto compacto  $K \subset \mathbb{R}$  y para todo  $x \in K$  existen dos constantes  $a, b > 0$  tales que:

$$|f(x)| \leq a|x| + b.$$

**Definición 1.4.11 [Condición de Lipschitz]:** Sea  $f : \mathbb{R} \mapsto \mathbb{R}$  una función, decimos que  $f$  satisface la condición de Lipschitz (o  $f$  es Lipschitz continua) si para cualesquiera  $x, y \in \mathbb{R}$  existe una constante  $L > 0$  tal que:

$$|f(x) - f(y)| \leq L|x - y|.$$

Por otro lado, decimos que  $f$  satisface la condición de Lipschitz local si existe un subconjunto compacto  $K \subset \mathbb{R}$  y para cualesquiera  $x, y \in K$  existe una constante  $L > 0$  tal que:

$$|f(x) - f(y)| \leq L|x - y|.$$

**Teorema 1.4.12 [Criterio de continuidad de Kolmogorov]:** Sea  $X = \{X_t\}_{t \geq 0}$  un proceso estocástico real tal que existen tres constantes positivas  $\alpha, \beta, C$  tales que:

$$\mathbb{E}[|X_t - X_s|^\alpha] \leq C|t - s|^{1+\beta},$$

para cualesquiera  $s, t \geq 0$ . Entonces,  $X$  posee una modificación continua  $\tilde{X}$  la cual, con probabilidad 1, es localmente Hölder continua de exponente  $\gamma$  tal que  $0 < \gamma < \frac{\beta}{\alpha}$  (véase **Definición 1.4.3**).

(La demostración y definiciones de este resultado puede encontrarse en [3]).

**Lema 1.4.13 [Principio del máximo]:** Sea  $D \subset \mathbb{R}$  un conjunto abierto y  $\mathbb{S} \subset \mathbb{N}$ . Supóngase que  $D$  es conexo y sea  $f \geq 0$  una función  $\mathcal{L}$ -armónica en  $D \times \mathbb{S}$ . Entonces  $f$  es estrictamente positiva en  $D \times \mathbb{S}$  o es idénticamente cero.

(La demostración y definiciones de este resultado pueden encontrarse en [18]).

**Lema 1.4.14 [Lema de Fatou]:** Sean  $f \in M(X, \mathbb{X})$  una función real-valuada no negativa y medible en un espacio  $X$  y  $\{\mu_n\}_{n \in \mathbb{N}}$  una sucesión de medidas en  $\mathbb{X}$ . Si  $\liminf_{n \rightarrow \infty} \mu_n(A) \geq \mu(A)$  para cualquier  $A \in \mathbb{X}$  para alguna medida  $\mu$ , entonces:

$$\liminf_{n \rightarrow \infty} \int_X f(x) d\mu_n(x) \geq \int_X f(x) d\mu(x).$$

(La demostración de este resultado puede encontrarse en [24]).

**Teorema 1.4.15 [Teorema de convergencia dominada]:** Supóngase que se satisfacen las condiciones del Lema 1.4.14, con la única diferencia de que  $\mu_n(A) \rightarrow \mu(A)$  para cualquier  $A \in \mathbb{X}$  y que  $\mu_n \leq \nu$  para algunas medidas  $\mu, \nu$  tal que  $\int_X \nu < \infty$ , entonces:

$$\lim_{n \rightarrow \infty} \int_X f(x) d\mu_n(x) = \int_X f(x) d\mu(x).$$

(La demostración de este resultado puede encontrarse en [24]).

**Teorema 1.4.16 [Fórmula de Dynkin]:** Sea  $X = \{X_t\}_{t \geq 0}$  un proceso de difusión (véanse definiciones 1.2.1 y 1.2.2) que satisface la siguiente ecuación diferencial estocástica:

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dw(t),$$

donde  $\mu(x), \sigma(x), w(\cdot)$  son el coeficiente de deriva, difusión y un movimiento Browniano regular respectivamente.

Sean  $x \in \mathbb{S}, \tau$  un tiempo de paro (Definición 1.2.8) de forma que  $\mathbb{E}[\tau|X_0] < \infty$  y  $f \in \mathcal{C}^2(\mathbb{S})$ . Entonces se satisface que:

$$\mathbb{E}[f(X_\tau)|X_0 = x] = f(x) + \mathbb{E}\left[\int_0^\tau \mathcal{L}f(X_s)ds|X_0 = x\right],$$

donde  $\mathcal{L}$  es el operador infinitesimal de  $X$  (véase Definición 1.2.5).

(La demostración de este resultado puede encontrarse en [14]).

**Definición 1.4.17 [Martingala]:** Sea  $(\Omega, \mathcal{F}, P)$  un espacio de probabilidad. Sea  $\{\mathcal{F}_t\}$  una filtración (véase Definición 1.1.2 para el caso  $t \in \mathbb{N}$ ). Sea  $\{X_n, n \geq 1\}$  una sucesión de variables aleatorias con espacio de estados  $\Omega$  que satisfacen las siguientes condiciones:

- a)  $X_n$  es medible con respecto a  $\mathcal{F}_n$ ,
- b)  $X_n$  es integrable,
- c) Si  $m < n$  y  $A$  es cualquier conjunto contenido en  $\mathcal{F}_m$ , entonces:

$$\int_A X_n dP = \int_A X_m dP. \tag{26}$$

Entonces diremos que  $\{X_n, n \geq 1\}$  es una martingala con respecto a  $\{\mathcal{F}_t\}_{t \in \mathbb{N}}$ . Ahora, para el caso donde las variables aleatorias sean reales y si reemplazamos '=' por ' $\leq$ ' o por ' $\geq$ ' en (26), a la sucesión de variables aleatorias se le llamará supermartingala o submartingala, respectivamente con respecto a  $\{\mathcal{F}_t\}_{t \in \mathbb{N}}$ .

**Lema 1.4.18 [Fórmula generalizada de Itô]:** Sea  $X(t) = (X_1(t), \dots, X_d(t)) = X = (X_1, \dots, X_d)$  una semi-martingala  $d$ -dimensional (es decir puede ser una submartingala o supermartingala) a tiempo continuo que toma valores en un subconjunto abierto  $U \subset \mathbb{R}^d$ . Entonces para cualquier función doblemente diferenciable  $f : U \mapsto \mathbb{R}$ ,  $f(X)$  es una semi-martingala y

$$df(X) = D_i f(X) dX_i + \frac{1}{2} D_{ij} f(X) d[X_i, X_j],$$

donde  $D_i$  es la derivada de orden 1 con respecto a  $X_i$  y  $D_{ij}$  es la derivada de segundo orden con respecto a  $X_i$  seguido de  $X_j$ .

Para ser más claros, reescribiendo la expresión anterior en su forma integral y explícitamente sumando sobre todos los índices, tendremos que:

$$f(X) = f(X(0)) + \sum_{i=1}^d \int D_i f(X) dX_i + \frac{1}{2} \sum_{i,j=1}^d \int D_{ij} f(X) d[X_i, X_j],$$

donde  $X(0)$  es el valor de la semi-martingala  $X$  a tiempo  $t = 0$ .

(La demostración de este resultado puede encontrarse en [7]).

**Definición 1.4.19 [Martingala local]:** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $\mathcal{F} = \{\mathcal{F}_t\}$  una filtración,  $X = \{X_t : t \geq 0\}$  cuyo espacio de estados es  $\Omega$ . Diremos que  $X$  es una martingala local con respecto a  $\mathcal{F}$  si existe una sucesión de tiempos de paro  $\tau_k$  tales que:

- i) la sucesión  $\tau_k$  es casi seguramente creciente:  $\mathbb{P}[\tau_k < \tau_{k+1}] = 1$ .
- ii) la sucesión  $\tau_k$  es casi seguramente divergente:  $\mathbb{P} \left[ \lim_{k \rightarrow \infty} \tau_k = \infty \right] = 1$ .
- iii) el 'proceso de paro'  $X_t^{\tau_k} := X_{\min\{t, \tau_k\}}$  es una martingala con respecto a  $\mathcal{F}$  para todo  $k$ .

**Teorema 1.4.20 [Desigualdad de Burkholder-Davis-Gundy]:** Para cualquier  $p \in [1, \infty)$  existen dos constantes positivas  $c_p, C_p$  tales que, para toda martingala local  $X$  tal que  $X_0 = 0$  y tiempos de paro  $\tau$ , se satisface la siguiente desigualdad:

$$c_p \mathbb{E} \left[ \left( \sum_{i \geq 1} (X_i - X_{i-1})^2 \right)^{\frac{p}{2}} \right] \leq \mathbb{E} \left[ \left( \sup_{n \geq 0} |X_n| \right)^p \right] \leq C_p \mathbb{E} \left[ \left( \sum_{i \geq 1} (X_i - X_{i-1})^2 \right)^{\frac{p}{2}} \right].$$

(La demostración de este resultado puede encontrarse en [9]).

**Definición 1.4.21 [Medida de martingala]:** Sea  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$  un espacio de probabilidad filtrado. Considérese  $\mathbb{B}(\mathbb{R}^r)$  la  $\sigma$ -álgebra de Borel en  $\mathbb{R}^r$ . Sea  $m(\cdot, \cdot) : \mathbb{R}^r \times \mathbb{R}_+ \mapsto \mathbb{R}$ . Diremos que  $m$  es una medida de martingala con respecto a  $(\{\mathcal{F}_t\}, \mathbb{P})$  si se satisfacen las siguientes propiedades:

1. Para cada  $A \in \mathbb{B}(\mathbb{R}^r)$ ,  $m(A, \cdot)$  es una martingala cuadrada integrable con respecto a  $(\{\mathcal{F}_t\}, \mathbb{P})$

$$\text{y } m(A, 0) = 0.$$

2. Para cualesquiera  $A, B \in \mathbb{B}(\mathbb{R}^r)$  tales que  $A \cap B = \emptyset$ , tendremos que

$m(A \cup B, t) = m(A, t) + m(B, t)$  casi seguramente en probabilidad para todo  $t > 0$ .

3. Para todo  $t > 0$ ,  $m(\cdot, t)$  satisface que existe una sucesión no decreciente  $\{E_n\}_{n \in \mathbb{N}}$

de subconjuntos de Borel de  $\mathbb{R}^r$  tales que  $\bigcup_{n=1}^{\infty} E_n = \mathbb{R}^r$  de manera que:

a) para todo  $t > 0$ ,  $\sup_{A \in \epsilon_n} \mathbb{E} [m(A, t)^2] < \infty$ ,  $\epsilon_n = \mathbb{B}(E_n)$ ,

b) para todo  $t > 0$ ,  $\mathbb{E} [m(A_j, t)^2] \rightarrow 0$ , para cualquier sucesión  $A_j$  contenida en  $\epsilon_n$  que

decrece a  $\emptyset$ .



## 2 Capítulo II: Construcción y desarrollo de procesos de difusión cambiante

En este capítulo desarrollaremos el concepto de *procesos de difusión cambiante*, desde su definición, funciones y/u operadores asociados hasta el análisis de propiedades y resultados fundamentales de esta clase de procesos. En esta última parte se adaptan y vuelven a definir algunos conceptos que ya habían sido definidos en la Sección 1.2, desde la *función de distribución de probabilidad del proceso* (véase **Definición 1.2.3**), pasando por el *operador infinitesimal* (véase **Definición 1.2.6**), siguiendo con *recurrencia* (véase **Definición 1.2.11**), hasta la *propiedad de Feller* (véase **Definición 1.2.13**). Además de los conceptos anteriormente mencionados, en este capítulo se incluyen definiciones y resultados con respecto a la *ergodicidad* de un proceso de difusión cambiante. Cabe mencionar que a lo largo de este capítulo se desarrollan ejemplos teóricos y gráficos de las propiedades que satisfacen esta clase de procesos. Estos ejemplos se programaron y simularon en el lenguaje de programación **R**.

### 2.1 Procesos de difusión cambiante y propiedades importantes

Vamos a trabajar en un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$  dotado de una *filtración*  $\{\mathcal{F}_t\}$  para todo  $t \geq 0$ . Sea  $(X(t), Y(t))$  un proceso de Markov homogéneo que toma valores en  $\mathbb{S} \times S$ , donde  $\mathbb{S} = (a, b) \subseteq \mathbb{R}$  (usualmente se supone que  $\mathbb{S} \subset \mathbb{R}^r : r \in \mathbb{N}$ ) y  $S = \{1, \dots, N\} : N \leq \infty$  es un conjunto finito o numerable (posiblemente  $\mathbb{N}_0$ ).

Supóngase que  $Y(\cdot)$  es un proceso estocástico con trayectorias continuas por la derecha (en otras palabras un proceso de saltos), el cual toma valores en  $S$  y que posee un generador  $Q(x)$  el cual actúa sobre funciones vectoriales  $f(x, \cdot)$ , de forma tal que:

$$\begin{aligned} Q(x)f(x, \cdot)(i) &= \sum_{j=1}^N q_{ij}(x)f(x, j) \\ &= \sum_{j \in S} q_{ij}(x)(f(x, j) - f(x, i)), \text{ para cada } i \in S, \end{aligned} \tag{27}$$

de forma que para cada  $i \in S$  se satisface que  $\sum_{j \in S} q_{ij}(x) = 0$ .

**Nota:** más adelante en esta sección vamos a redefinir a  $Q(x)$ , esta vez lo nombraremos *matriz de intensidad de transición* y desarrollaremos algunos resultados importantes con respecto a este generador.

Ahora, sea  $w(\cdot)$  un movimiento Browniano regular en  $\mathbb{R}^r$  definido en el espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ . Supóngase la existencia de funciones  $b(\cdot, \cdot) : \mathbb{S} \times S \rightarrow \mathbb{R}$  y  $\sigma(\cdot, \cdot) : \mathbb{S} \times S \rightarrow \mathbb{R}^r$ . Entonces, diremos que  $(X(t), Y(t))$  es un *proceso de difusión cambiante* si:

$$\begin{aligned} dX(t) &= b(X(t), Y(t))dt + \sigma(X(t), Y(t))dw(t), \\ (X(0), Y(0)) &= (x, \alpha) : \alpha \in S, x \in \mathbb{S}, \end{aligned} \tag{28}$$

y para cualesquiera  $i \neq j \in S$ ,

$$\mathbb{P}[Y(t + \Delta) = j | Y(t) = i, X(s), Y(s), s \leq t] = q_{ij}(X(t))\Delta + o(\Delta). \tag{29}$$

**Definición 2.1.1:** Las probabilidades de transición son descritas a través de la función matricial  $P(t; x, A)$ , cuyas entradas  $(i, j)$  son  $\mathbb{P}[X(t) \in A, Y(t) = j | X(0) = x, Y(0) = i]$  donde  $t \geq 0, x \in \mathbb{S}$  y  $A \subseteq \mathbb{S}$  es un conjunto Borel de  $\mathbb{R}$ .

Bajo condiciones generalizadas, se puede demostrar que el proceso puede caracterizarse a través de una familia de procesos de difusión  $\{X_i(t) | i \in S\}$ , donde cada proceso  $X_i(t)$  tiene un tiempo finito de vida aleatorio

de tal forma que la componente continua  $X(t)$  se comporta como el proceso  $X_i(t)$  durante la estadía del componente discreto  $Y(t)$  en el estado  $i$ . La duración de  $Y$  en  $i$  coincide con el tiempo de vida asociado al proceso  $X_i$ .

**Definición 2.1.2:** Para cualquier función vectorial Borel acotada  $f(x) = [f_i(x)]_{i \in S}$  se define el operador de transición  $T_t f$  como:

$$T_t f(x) = \mathbb{E} [f_{Y(t)}(X(t)) | X(0) = x, Y(0) = i] = \int_S \sum_{j \in S} P_{ij}(t; x, dy) f_j(y) = i\text{-ésima componente del vector } \int_S P(t; x, dy) f(y). \quad (30)$$

Asumimos que la familia de funciones  $P(t; x, A)$  satisface la ecuación de Chapman-Kolmogorov:

$$P_{ij}(s+t; x, A) = \int_S \sum_{h \in S} P_{ih}(s; x, dy) P_{hj}(t; y, A), \quad (31)$$

o en forma matricial:

$$P(s+t; x, A) = \int_S P(s; x, dy) P(t; y, A). \quad (32)$$

Considérese la **Definición 1.4.1**. Ahora, supóngase que  $(X(t), Y(t))$  comienza en  $(x, i)$  a tiempo  $t = 0$ . Al par anterior le asociamos un proceso de difusión  $X_i(t)$  que comienza en el punto  $x$  y un generador  $\mathcal{L}_i f$  de la forma

$$(\mathcal{L}_i f)(x) = \frac{1}{2} \sigma^2(x, i) f''(x) + b(x, i) f'(x) + Q_i(x) f(x),$$

donde  $Q_i(x) < 0$  y  $-Q_i(x)$  representa la tasa de término del proceso  $X_i(t)$  en el punto  $x$ .  $Q_i(x)$  determina el tiempo de vida del proceso  $X_i(\cdot)$ . Al momento en que el proceso  $X_i(\cdot)$  muere, la componente  $Y(t)$  selecciona otro estado  $j$  con probabilidad  $\frac{-q_{ij}(x)}{q_{ii}(x)}$ ,  $j \neq i$ , donde  $x$  representa el punto del proceso  $X_i(\cdot)$  al momento de su término. Por ende las intensidades de transición entre estados son afectadas por el valor al momento de término del componente continuo.

Vamos a suponer la existencia de una función matricial  $Q(x)$ , de tal forma que

$$\lim_{h \rightarrow 0} \frac{1}{h} [P(h; x, S) - I_{N \times N}] = Q(x),$$

misma a la que vamos a llamar ‘matriz de intensidad de transición de  $(X(t), Y(t))$ ’. Al momento de transición,  $(X(t), Y(t))$  está en el estado  $(x, j)$ . El procedimiento del estado  $i$  se repite para el estado  $j$ . A dicho estado le asociamos un proceso de difusión  $X_j(t)$  que comienza en  $x$ , el cual tiene asociado un tiempo de vida cuya distribución es determinada por medio de una tasa de término  $-Q_j(x)$ .  $Y_j(t)$  permanece en  $j$  por lo que dura el tiempo de vida de  $X_j(t)$ . Luego selecciona otro estado al momento de que  $X_j(\cdot)$  muere, y así sucesivamente (véase la **Definición 1.4.1**, en este caso la diferencia entre la matriz  $Q$  y  $Q(x)$  se encuentra en el hecho de que para la cadena de Markov no hay un cambio en cuanto a las entradas de la matriz, mientras que para el proceso de difusión cambiante la intensidad de transición del componente discreto  $Y(\cdot)$  dependerá en todo momento del estado del componente continuo  $X(\cdot)$ ).

Ahora, de regreso a la matriz de densidad de transición  $P(t; x, A)$ , para facilitar el análisis, vamos a suponer que cada entrada  $P_{ij}(t; x, dy)$  posee una densidad  $P_{ij}(t; x, y)$  que satisface

$$P_{ij}(t; x, A) = \int_A P_{ij}(t; x, y) dy. \quad (33)$$

Tenemos que para todo conjunto Borel  $A \subset \mathbb{S}$

$$\left( \int_A P(t; x, y) dy \right) e_N \leq e_N, \quad (34)$$

donde  $e_N = (1, 1, \dots, 1)^T$ . De forma tal que la ecuación de Chapman-Kolmogorov ahora se reescribe como:

$$P(s+t; x, y) = \int_{\mathbb{S}} P(s; x, z) P(t; z, y) dz. \quad (35)$$

Al proceso  $(X(t), Y(t))$  le asociamos un operador infinitesimal  $\mathcal{L}$ . Para todo  $i \in S$  y cualquier  $f(\cdot, i) \in \mathcal{C}^2$ , donde  $\mathcal{C}^2$  denota a la clase de funciones cuyas derivadas parciales con respecto a la variable  $x$  hasta segundo orden con continuas, tendremos:

$$\begin{aligned} \mathcal{L}f(x, i) &= \text{tr}(\nabla^2 f(x, i) a(x, i)) + \nabla f^t(x, i) b(x, i) + Q(x) f(x, \cdot)(i) \\ &= \frac{1}{2} \sum_{j, k=1}^r a_{jk}(x, i) \frac{\partial^2 f(x, i)}{\partial x_j \partial x_k} + \sum_{k=1}^r b_k(x, i) \frac{\partial f(x, i)}{\partial x_k} + Q(x) f(x, \cdot)(i), \end{aligned} \quad (36)$$

donde  $\nabla f(x, i)$  y  $\nabla^2 f(x, i)$  son el gradiente y el Hessiano de la función  $f(x, i)$  con respecto a  $x$ , respectivamente y  $\mathbb{S} \subset \mathbb{R}^r$ . Además:

$$\begin{aligned} Q(x) f(x, \cdot)(i) &= \sum_{j=1}^N q_{ij}(x) f(x, j), \\ a(x, i) &= (a_{jk}(x, i)) = \sigma(x, i) \sigma^*(x, i) \in \mathbb{R}^{r \times r}. \end{aligned}$$

Para el caso unidimensional  $r = 1$  se tiene que el operador infinitesimal es el siguiente:

$$\mathcal{L}f(x, i) = \frac{1}{2} \sigma^2(x, i) \frac{d^2}{dx^2} f(x, i) + b(x, i) \frac{d}{dx} f(x, i) + Q(x) f(x, \cdot)(i).$$

**Nótese** que la evolución del componente discreto  $Y(\cdot)$  puede y es representada a través de una integral estocástica con respecto a una medida de Poisson aleatoria (véase [36]). En efecto, para  $x \in \mathbb{S}$  e  $i \neq j \in S$ , sean  $\Delta_{ij}(x)$  los intervalos reales consecutivos (con respecto a un orden lexicográfico en  $S \times S$ ) cerrados por la izquierda y abiertos por la derecha, cada uno de longitud  $q_{ij}(x)$ . Definimos la función:

$$h(x, i, z) = \sum_{j=1}^N (j - i) \mathbb{1}_{z \in \Delta_{ij}(x)}.$$

Es decir, ocupando la partición  $\{\Delta_{ij}(x) : i, j \in S\}$ , si  $z \in \Delta_{ij}(x)$ ,  $h(x, i, z) = \sum_{j=1}^N j - i$ , de otra forma  $h(x, i, z) = 0$ . De esta forma (28) es equivalente a:

$$dY(t) = \int_{\mathbb{R}} h(X(t), Y(t-), z) p(dt, dz),$$

donde  $p(dt, dz)$  es una medida aleatoria Poisson de intensidad  $dt \times m(dz)$ ,  $Y(t_a-)$  es el límite por la izquierda del componente  $Y(\cdot)$  al aproximar  $t$  a  $t_a$  por ‘abajo’, o bien  $t \uparrow t_a$  y  $m$  es la medida de Lebesgue en  $\mathbb{R}$ . Cabe resaltar que la medida aleatoria Poisson  $p(\cdot, \cdot)$  es independiente al movimiento Browniano  $w(\cdot)$ .

**Teorema 2.1.3:** *Ocupando  $\mathcal{L}$  definido en (36), para todo  $f(\cdot, i) \in \mathcal{C}^2$ ,  $i \in S$  se satisface que:*

$$f(X(t), Y(t)) - f(X(0), Y(0)) = \int_0^t \mathcal{L}f(X(s), Y(s)) ds + M_1(t) + M_2(t), \quad (37)$$

donde:

$$M_1(t) = \int_0^t \langle \nabla f(X(s), Y(s)), \sigma(X(s), Y(s)) dw(s) \rangle,$$

$$M_2(t) = \int_0^t \int_{\mathbb{R}} [f(X(s), Y(0) + h(X(s), Y(s), z)) - f(X(s), Y(s))] \mu(ds, dz),$$

$\mu(ds, dz) = p(ds, dz) - ds \times m(dz)$  es una medida de martingala (véase **Definición 1.4.21**) y las funciones  $M_1(t), M_2(t)$  son semi-martingalas (véase **Definición 1.4.17**).

(La demostración de este resultado puede encontrarse en [36], [7]).

**Nótese** que ocupando el anterior resultado, tendremos que:

$$M_f(t) = f(X(t), Y(t)) - f(X(0), Y(0)) - \int_0^t \mathcal{L}f(X(s), Y(s)) ds$$

es una martingala local (véase **Definición 1.4.19**). Si para cualquier  $i \in S$ ,  $f(x, i) \in \mathcal{C}_b^2$  (la clase de funciones de derivadas parciales continuas y acotadas hasta segundo orden con respecto a  $x$ ), o  $f(x, i) \in \mathcal{C}_0^2$  (funciones de la clase  $\mathcal{C}^2$  de soporte compacto), entonces  $M_f(t)$  se convierte en martingala (véase **Definición 1.4.17**).

**Otra consecuencia** de la fórmula generalizada de Itô (37), (véase **Lema 1.4.18** para el caso de procesos de difusión) es que la *Fórmula de Dynkin* (véase **Teorema 1.4.16**) se satisface, la cual se enuncia en el siguiente teorema:

**Teorema 2.1.4:** Sean  $f(\cdot, i) \in \mathcal{C}^2$  para algún  $i \in S$ , y  $\tau_1, \tau_2$  tiempos de paro acotados de forma tal que  $0 \leq \tau_1 \leq \tau_2$  casi seguramente. Si  $f(X(t), Y(t))$  y  $\mathcal{L}f(X(t), Y(t))$  están acotados para cualquier  $t \in [\tau_1, \tau_2]$  con probabilidad 1, entonces:

$$\mathbb{E}[f(X(\tau_2), Y(\tau_2))] = \mathbb{E}[f(X(\tau_1), Y(\tau_1))] + \mathbb{E}\left[\int_{\tau_1}^{\tau_2} \mathcal{L}f(X(s), Y(s)) ds\right].$$

(La demostración de este resultado puede encontrarse en [36], [25]).

**Definición 2.1.5:** Considérese la función  $Q(x) = (q_{ij}(x)) : \mathbb{R}^r \rightarrow \mathbb{R}^{N \times N}$ , decimos que  $Q$  satisface la propiedad  $q$  si para cualesquiera  $x \in \mathbb{R}^r$  e  $i, j \in S$ , la función  $q_{ij}(x)$  es Borel medible, uniformemente acotada,  $q_{ij}(x) \geq 0 : i \neq j$  y  $q_{ii}(x) = -\sum_{j \neq i} q_{ij}(x)$ .

**Teorema 2.1.6:** Sean  $x \in \mathbb{R}^r$ ,  $S = \{1, \dots, N\}$  y  $Q(x) = (q_{ij}(x))$  la función matricial de tamaño  $N \times N$  que satisface la propiedad  $q$ . Considérese el proceso de dos componentes  $(X(t), Y(t))$  definido al principio de la Sección 2.1, tal que  $(X(0), Y(0)) = (x, \alpha)$ . Supóngase que  $Q(\cdot) : \mathbb{R}^r \rightarrow \mathbb{R}^{N \times N}$  es una función continua y acotada, y que las funciones  $b(\cdot, \cdot)$  y  $\sigma(\cdot, \cdot)$  en (28) satisfacen lo siguiente:

$$|b(x, \alpha)| + |\sigma(x, \alpha)| \leq K(1 + |x|), \quad \alpha \in S. \quad (38)$$

Supongamos, para todo entero  $N \geq 1$  existe una constante positiva  $M_N$ , tal que para todo  $t \in [0, T]$ ,  $i \in S$  y cualesquiera  $x, y \in \mathbb{R}^r$  donde  $|x| \vee |y| \leq M_N$ , se satisface que:

$$|b(x, i) - b(y, i)| \vee |\sigma(x, i) - \sigma(y, i)| \leq M_N |x - y|. \quad (39)$$

Si todo lo anterior se cumple, entonces existe una solución única  $(X(t), Y(t))$  a la ecuación (28) con información inicial de forma que la evolución del proceso de salto esté especificada por (29).

(La demostración de este resultado puede encontrarse en [36], [25]).

**Proposición 2.1.7:** *Supóngase que se satisfacen las condiciones del Teorema 2.1.6. Sea  $T > 0$  fijo. Entonces, para cualquier constante positiva  $\gamma$  tenemos que:*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |X(t)|^\gamma \middle| (X(0), Y(0)) = (x, i) \right] \leq C < \infty, \quad (x, i) \in \mathbb{S} \times S,$$

donde la constante  $C$  depende de  $x, T, \gamma$ , i.e.  $C = C(x, T, \gamma)$ .

**Demostración:**

*Paso 1*

Dado que

$$X(t) = x + \int_0^t b(X(s), Y(s)) ds + \int_0^t \sigma(X(s), Y(s)) dw(s),$$

entonces para todo  $p \geq 2$ :

$$|X(t)|^p \leq 3^{p-1} \left[ |x|^p + \left( \left| \int_0^t b(X(s), Y(s)) ds \right| \right)^p + \left( \left| \int_0^t \sigma(X(s), Y(s)) dw(s) \right| \right)^p \right].$$

Ocupando la desigualdad de Hölder (**Proposición 1.4.4**) y la condición de crecimiento lineal (véase **Definición 1.4.10**) llegamos a que:

$$\begin{aligned} & \mathbb{E} \left[ \sup_{t \in [0, T]} \left( \left| \int_0^t b(X(s), Y(s)) ds \right| \right)^p \middle| (X(0), Y(0)) = (x, i) \right] \\ & \leq c_1(T, p) \int_0^{T_1} \mathbb{E}_{(x, i)} \left[ 1 + \sup_{0 \leq u \leq s} |X(s)|^p \right] ds, \end{aligned}$$

donde  $0 \leq T_1 \leq T$  y  $c_1(T, p)$  es una constante positiva que depende únicamente de  $T, p$  y la constante de linealidad en (38). De forma análoga, debido a las desigualdades de Burkholder-Davis-Gundy (véase página 70 de [28] y **Teorema 1.4.20**), de Hölder (véase **Proposición 1.4.4**) y la condición de crecimiento lineal (véase **Definición 1.4.10**) tenemos que:

$$\mathbb{E}_{(x, i)} \left[ \sup_{0 \leq t \leq T} \left| \int_0^t \sigma(X(s), Y(s)) dw(s) \right|^p \right] \leq c_2(T, p) \int_0^{T_1} \mathbb{E}_{(x, i)} \left[ 1 + \sup_{0 \leq u \leq s} |X(u)|^p \right] ds,$$

donde  $0 \leq T_1 \leq T$  y  $c_2(T, p)$  es una constante positiva que depende únicamente de  $T, p$  y la constante de linealidad en (38). Por lo tanto:

$$\begin{aligned} & \mathbb{E}_{(x, i)} \left[ 1 + \sup_{0 \leq t \leq s} |X(t)|^p \right] \\ & \leq c_3(x, p) + c_4(T, p) \int_0^{T_1} \mathbb{E}_{(x, i)} \left[ 1 + \sup_{0 \leq u \leq s} |X(u)|^p \right] ds, \end{aligned}$$

donde  $c_3(x, p) = 1 + 3^{p-1}|x|^p$ , y  $c_4(T, p) = 3^{p-1}|c_1(T, p) + c_2(T, p)|$ . Por lo que la desigualdad de Gronwall (véase **Proposición 1.4.6**) implica que:

$$\mathbb{E}_{(x,i)} \left[ 1 + \sup_{0 \leq t \leq s} |X(t)|^p \right] \leq c_3(x, p)e^{c_4(T,p)T_1} \leq c_3(x, p)e^{c_4(T,p)T} := C(x, T, p).$$

Tenemos que  $\mathbb{E}_{(x,i)} \left[ \sup_{0 \leq t \leq s} |X(t)|^p \right] \leq C(x, T, p)$ . Dado que  $T_1 \leq T$  es arbitrario, se sigue que:

$$\mathbb{E}_{(x,i)} \left[ \sup_{0 \leq t \leq s} |X(t)|^p \right] \leq C(x, T, p), \quad p \geq 2.$$

*Paso 2*

Nótese que  $\sup_{0 \leq t \leq T} |X(t)|^\gamma = (\sup_{0 \leq t \leq T} |X(t)|)^\gamma$  para todo  $\gamma > 0$ . Ocupando el *Paso 1* y la desigualdad de Hölder (**Proposición 1.4.4**), se tiene que para todo  $1 \leq p < 2$ ,

$$\mathbb{E}_{(x,i)} \left[ \sup_{0 \leq t \leq s} |X(t)|^p \right] \leq \left( \mathbb{E}_{(x,i)} \left[ \sup_{0 \leq t \leq s} |X(t)|^{2p} \right] \right)^{\frac{1}{2}} \leq C(x, T, p) < \infty.$$

*Paso 3*

Finalmente, si  $0 < p < 1$  tendremos que:

$$|X(t)|^p = |X(t)|^p \mathbb{1}_{\{|X(t)| < 1\}} + |X(t)|^p \mathbb{1}_{\{|X(t)| \geq 1\}} \leq 1 + |X(t)|^{1+p}.$$

Se sigue del *Paso 2* que:

$$\mathbb{E}_{(x,i)} \left[ \sup_{0 \leq t \leq s} |X(t)|^p \right] \leq \mathbb{E}_{(x,i)} \left[ 1 + \sup_{0 \leq t \leq s} |X(t)|^{p+1} \right] \leq C(x, T, p).$$

Con lo cual se concluye la demostración de la prueba.

Q.E.D.

**Proposición 2.1.8:** *El proceso  $(X(t), Y(t))$  es càdlàg. Es decir, las trayectorias de  $(X(t), Y(t))$  son continuas por la derecha y tiene límites por la izquierda.*

**Demostración:**

Es bien sabido que las trayectorias del componente discreto  $Y(\cdot)$  son continuas por la derecha y tiene límites por la izquierda. Sólo resta probar esas propiedades para el componente continuo  $X(\cdot)$ . Sea  $0 < t \leq T$ , donde  $T$  es un número positivo fijo y considérese:

$$X(t) - X(s) = \int_s^t b(X(u), Y(u))du + \int_s^t \sigma(X(u), Y(u))dw(u).$$

Ocupando la **Proposición 2.1.7** llegamos a que:

$$\mathbb{E}_{(x,i)} [|X(t) - X(s)|^4] \leq C|t - s|^2,$$

donde  $C$  es una constante que depende de  $T$ , la condición inicial  $x$ , la condición de crecimiento lineal (véase **Definición 1.4.10**) y la constante de Lipschitz (véase **Definición 1.4.11**) de los coeficientes  $b$  y  $\sigma$ . Por lo tanto, el resultado se sigue del criterio de continuidad de Kolmogorov (véase **Definición 1.4.12**).

Q.E.D.

Para finalizar esta sección se presentará una serie de simulaciones programadas en R para evidenciar algunos patrones de comportamiento de los procesos de difusión cambiante, en este caso la visualización de trayectorias.

**Ejemplo 2.1.9:** Sea  $N = 3$  y  $\mathbb{S} = \mathbb{R}$ , vamos a simular las trayectorias del proceso  $(X(t), Y(t))$  que satisface la ecuación diferencial estocástica:

$$\begin{cases} dX(t) = b(X(t), Y(t))dt + \sigma(X(t), Y(t))dw(t) : & t \in [0, +\infty), \\ (X(0), Y(0)) = (0, 1) : & (X(t), Y(t)) \in \mathbb{R} \times \{1, 2, 3\}. \end{cases} \quad (40)$$

donde

$$\sigma^2(x, i) = i^2, b(x, i) = -ix : (x, i) \in \mathbb{R} \times \{1, 2, 3\}.$$

Además,  $w(t)$  denota un movimiento Browniano estándar.

Las tres fases evolucionan de la misma manera que un conjunto de procesos de Orstein-Uhlenbeck (véase [8]) con parámetros distintos. Los coeficientes de difusión evolucionan con respecto al número de la fase  $Y(t)$  y los coeficientes de deriva dependen de manera más intensa del número de la fase  $Y(t)$ . Y la transición de las fases depende de la matriz

$$Q(x) = \begin{bmatrix} -\frac{2+x^2}{1+x^2} & \frac{5+3x^2}{6(1+x^2)} & \frac{7+3x^2}{6(1+x^2)} \\ 1 + \frac{x^2}{400} & -2 - \frac{x^2}{100} & 1 + \frac{3x^2}{400} \\ \frac{1}{2} + \frac{1}{2}e^{-x^2} & \frac{1}{2} + \frac{1}{2}e^{-x^2} & -1 - e^{-x^2} \end{bmatrix}.$$

Generamos aproximaciones equidistantes en el intervalo de tiempo  $[0, 3]$  de tamaño  $dt = 2^{-10}$  para el proceso  $(X(t), Y(t))$  en (40).

## Simulación de un Proceso de Orstein–Uhlenbeck bivariado

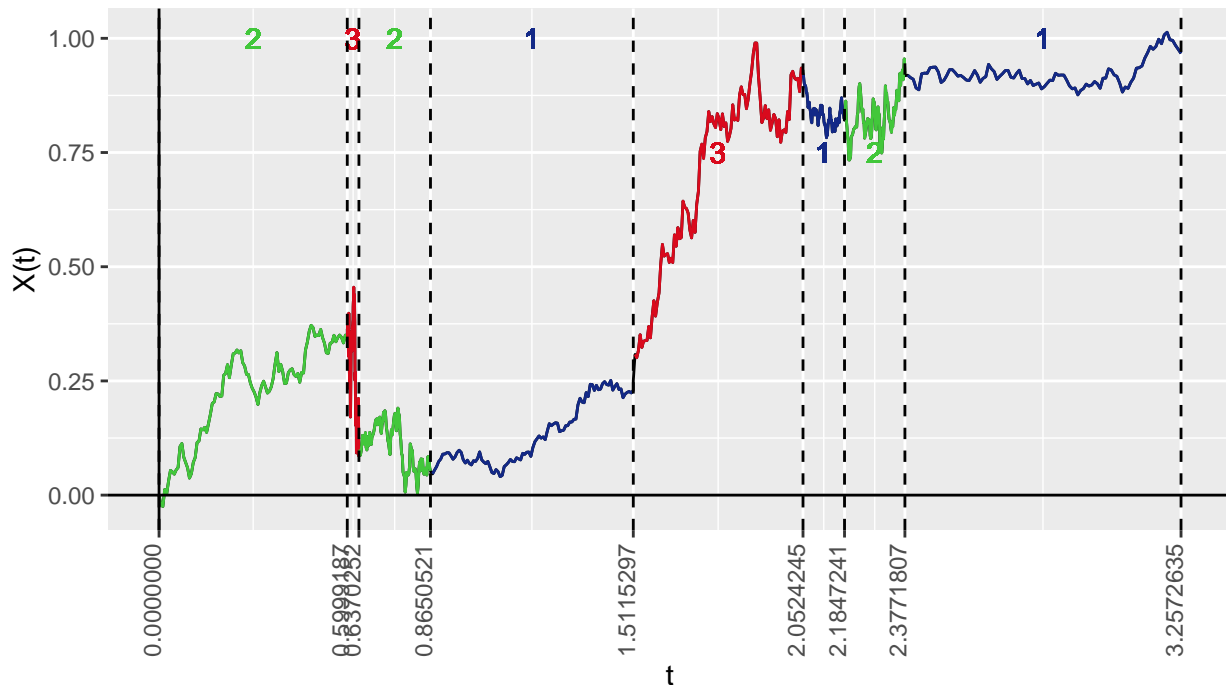


Figura 1: Trayectoria de un proceso de Markov bivariado de tres fases.

A continuación se muestra el código ocupado para realizar la simulación, la cual está basada en el Método de Euler [23] para simular un proceso estocástico. Cabe resaltar que este algoritmo se ocupa en todas las simulaciones subsecuentes, por lo cual en los próximos ejemplos omitiremos los códigos que se ocupan en las simulaciones.

```
# BLOQUE 1
# Cargamos el paquete 'spuRs' en nuestro entorno, ya que este permitirá
# trabajar con las cadenas de Markov a tiempo continuo.
library(spuRs)
# Cargamos el paquete 'ggplot2' con el objeto de facilitar la
# graficación de las trayectorias de los procesos.
library(ggplot2)
# Cargamos el paquete 'dplyr' para facilitar la notación y las
# llamadas a funciones.
library(dplyr)
# Cargamos el paquete 'sde' que facilita la labor de solucionar ecuaciones
# diferenciales estocásticas.
library(sde)
# La siguiente función crea la matriz de transición de fases y depende
# únicamente del valor del componente continuo X.
Q<-function(x){
  q11<-(-(2+x**2)/(1+x**2))
  q12<-((5+3*x**2)/(6+6*x**2))
  q13<-((7+3*x**2)/(6+6*x**2))
  q21<-1+(x**2/400)
  q22<-(-2-(x**2/100))
  q23<-1+(3*x**2/400)
```



```

q31<-(1/2)+(1/2)*exp(-(x**2))
q32<-(1/2)+(1/2)*exp(-(x**2))
q33<--1-exp(-(x**2))
q<-matrix(c(q11,q12,q13,q21,q22,q23,q31,q32,q33),
          nrow = 3, ncol = 3, byrow = TRUE)
return(q)
}
# Ocuparemos la función 'CMCSimulation' para generar la cadena de Markov Y(t)
# a tiempo continuo con matriz de intensidad Q(X(t)).

# NOTA IMPORTANTE: la función 'CMCSimulation' representará las fases
# como 0,1,2 y no como 1,2,3.

# La siguiente función simula el proceso completo al general la cadena de Markov,
# y al resolver n sistemas de ecuaciones diferenciales estocásticas,
# (n representa la cantidad de cambios de fase), se unen los resultados
# de cada solución y se grafica.

SDP1<-function(X0,Y0,T){
  set.seed(112)
  CMY<-CMCSimulation(Q(X0),Y0,T)
  I<-CMY[,1]+1
  Xfin<-c()
  Ind<-c()
  Tempo<-c()
  for (j in 2:dim(CMY)[1]) {
    if (I[j-1]==1){
      sigma<-expression(1)
      tau<-expression(-x)
    }
    else if (I[j-1]==2){
      sigma<-expression(2)
      tau<-expression(-2*x)
    }
    else {
      sigma<-expression(3)
      tau<-expression(-3*x)
    }
    # e=I[j-1]
    # sigma<-expression(e)
    # tau<-expression(-e*x)
    X<-sde.sim(CMY[,2][j-1],CMY[,2][j],X0,delta=0.0001,drift=tau,sigma=sigma)
    Xn<-as.vector(X)
    Xfin<-c(Xfin,Xn[-length(Xn)])
    X0<-tail(Xn,n=1)
    Ind<-c(Ind,rep(I[j-1],length(Xn)-1))
    Tempo<-c(Tempo,seq(CMY[,2][j-1],CMY[,2][j],length.out = length(Xn)-1))
  }
  return(list(Xfin,CMY,Ind,Tempo))
}

```

Nótese que la trayectoria simulada del proceso se define como una función a trozos que siempre es continua

para todo  $t > 0$ . A diferencia de un proceso de difusión regular este proceso puede moverse a través de  $N$  fases distintas con distintos coeficientes de difusión y deriva en general. Las transiciones a través de las fases y el tiempo que pasa entre cada fase depende de la posición  $X(t)$  y estos se obtienen a través de la matriz  $Q(x)$ .

En la *Figura 1* puede observarse la evolución del proceso en cada fase, al iniciar en  $X(0) = 0, Y(0) = 2$ . Nótese que la intensidad del proceso aumenta si el valor de  $Y(t)$  lo hace.

**Ejemplo 2.1.10:** *El siguiente ejemplo está enfocado en mostrar la diferencia entre los procesos de difusión simples y los procesos de difusión cambiante.*

Sea  $N = 2$  y  $\mathbb{S} = \mathbb{R}$ , vamos a simular las trayectorias del proceso  $(X_1(t), Y(t))$  que satisface la ecuación diferencial estocástica:

$$\begin{cases} dX_1(t) = b(X_1(t), Y(t))dt + \sigma(X_1(t), Y(t))dB(t) : & t \in [0, +\infty), \\ (X_1(0), Y(0)) = (0, 1) : & (X_1(t), Y(t)) \in \mathbb{R} \times \{1, 2\}. \end{cases} \quad (41)$$

donde

$$b(x, i) = \begin{cases} -0.2x & \text{si } i = 1 \\ 0.11x & \text{si } i = 2 \end{cases}, \quad \sigma(x, i) = \begin{cases} x & \text{si } i = 1 \\ 0.2x & \text{si } i = 2 \end{cases} : (x, i) \in \mathbb{R} \times \{1, 2\}.$$

La transición de las fases depende de la matriz

$$Q(x) = \begin{bmatrix} -5 \cos(x)^2 & 5 \cos(x)^2 \\ 10 \cos(x)^2 & -10 \cos(x)^2 \end{bmatrix}.$$

De manera similar, simularemos las trayectorias del proceso de difusión simple  $X_2(t)$  que satisface la ecuación diferencial estocástica:

$$\begin{cases} dX_2(t) = 0.11X_2(t)dt + 0.2X_2(t)dB(t) : & t \in [0, +\infty), \\ (X_1(0), Y(0)) = (3, 1) : & (X_1(t), Y(t)) \in \mathbb{R} \times \{1, 2\}. \end{cases} \quad (42)$$

Generemos aproximaciones equidistantes en el intervalo de tiempo  $[0, 3]$  de tamaño  $dt = 2^{-10}$  para los procesos  $(X_1(t), Y(t))$  y  $X_2(t)$  en (41) y (42), respectivamente.

## Diferencia entre un Proceso de Difusión Simple y Cambiante 1

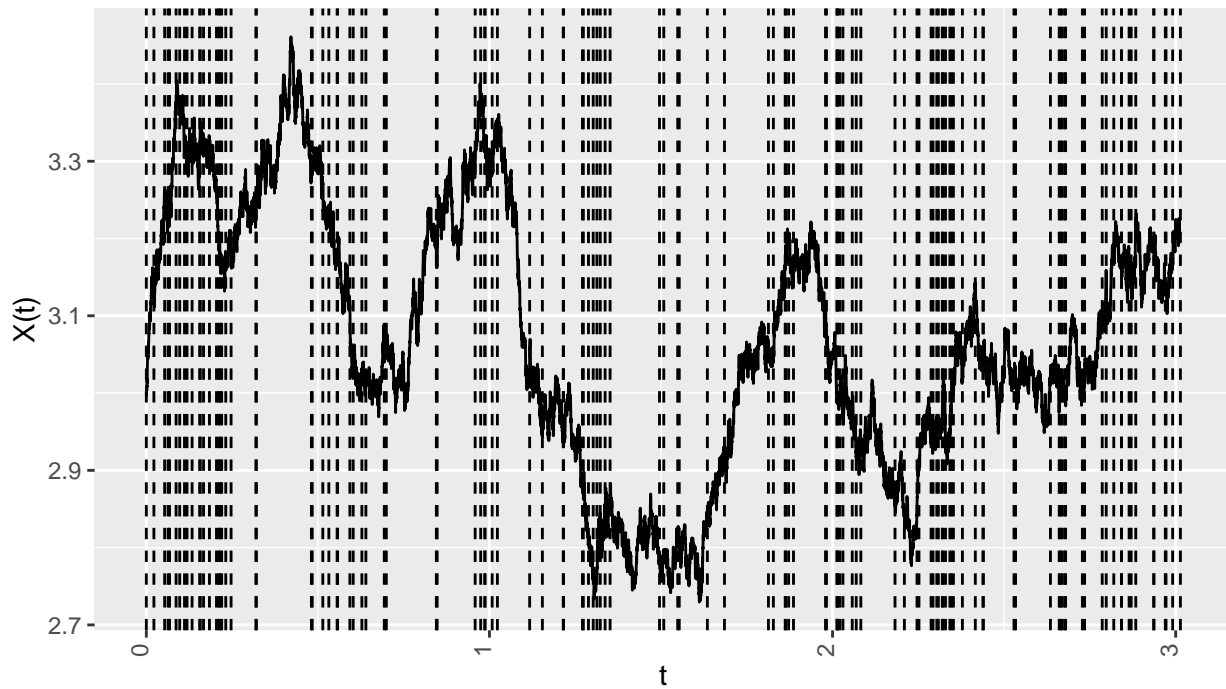


Figura 2.1: Trayectoria de un proceso de Markov bivariado de dos fases (41).

## Diferencia entre un Proceso de Difusión Simple y Cambiante 2

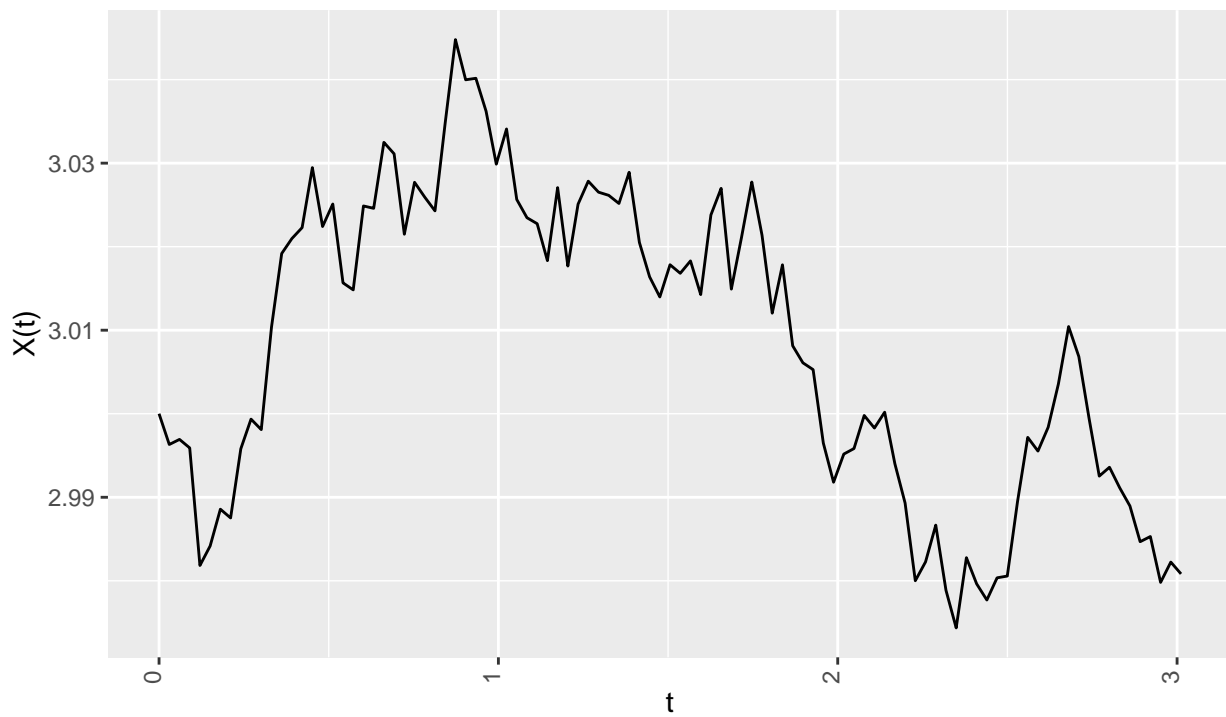


Figura 2.2: Trayectoria de un proceso de Markov (42).

El primer gráfico corresponde a la simulación del proceso (41), las líneas discontinuas paralelas al eje  $y$  representan el cambio de fase. El segundo gráfico corresponde a la simulación del proceso (42), debido a que este es un proceso de difusión simple entonces no habrá cambio en los coeficientes de deriva y difusión, por obvias razones no se incluyen líneas discontinuas como en la primera gráfica. Nótese que el primer proceso presenta una oscilación de valores mucho mayor a la del segundo proceso, es decir, mientras el componente continuo del primer proceso alcanza valores mayores 3.3 y menores a 2.9, el componente continuo del segundo proceso no alcanza valores mayores a 3.05 o menores a 2.97.

**Ejemplo 2.1.11:** En el siguiente ejemplo, el proceso tiene la particularidad de operar con coeficientes de deriva no lineales. Sea  $N = 2$  y  $S = \mathbb{R}$ , vamos a simular las trayectorias del proceso  $(X(t), Y(t))$  que satisface la ecuación diferencial estocástica:

$$\begin{cases} dX(t) = b(X(t), Y(t))dt + \sigma(X(t), Y(t))dB(t) : & t \in [0, +\infty), \\ (X(0), Y(0)) = (5, 1) : & (X(t), Y(t)) \in \mathbb{R} \times \{1, 2\}. \end{cases} \quad (43)$$

donde

$$b(x, i) = \begin{cases} -(2 + \sin(x)) & \text{si } i = 1 \\ -(1 + \sin(x) \cos(x)) & \text{si } i = 2 \end{cases}, \quad \sigma(x, i) = \begin{cases} x & \text{si } i = 1 \\ 0.2x & \text{si } i = 2 \end{cases} : (x, i) \in \mathbb{R} \times \{1, 2\}.$$

La transición de las fases depende de la matriz

$$Q(x) = \begin{bmatrix} -5 \cos(x)^2 & 5 \cos(x)^2 \\ 10 \cos(x)^2 & -10 \cos(x)^2 \end{bmatrix}.$$

Generemos aproximaciones equidistantes en el intervalo de tiempo  $[0, 5]$  de tamaño  $dt = 2^{-10}$  para el proceso  $(X(t), Y(t))$  descrito en (43).

### Ejemplo de una difusión cambiante con coeficientes de difusión no lineales

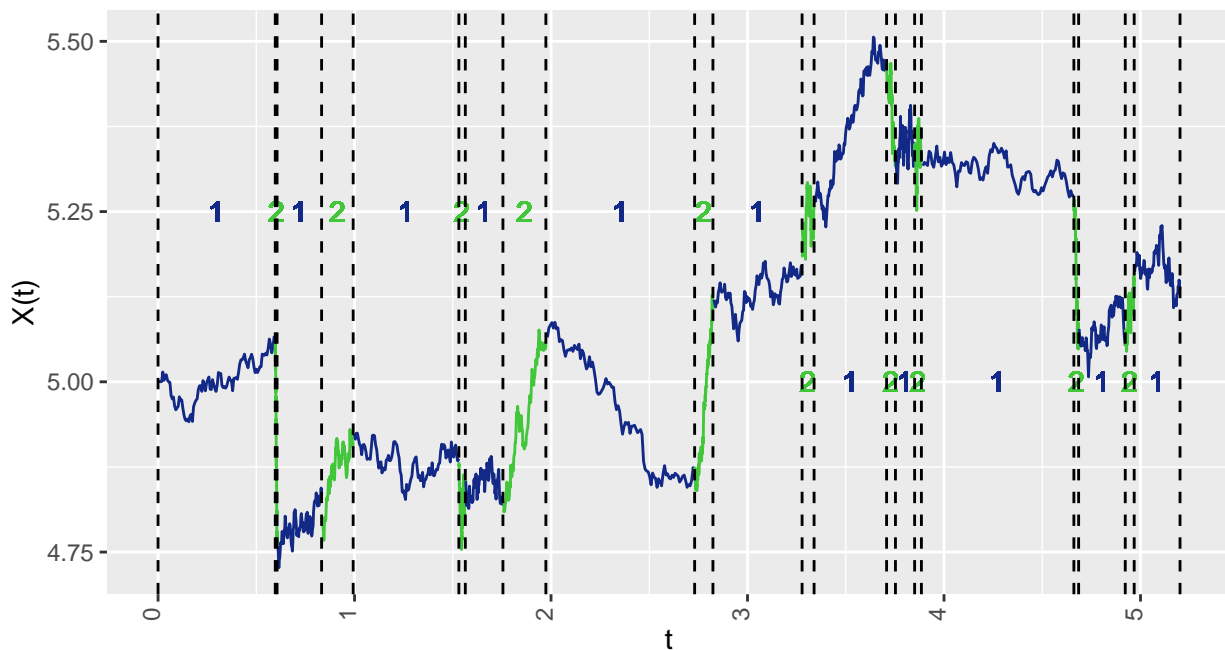


Figura 3: Trayectoria de un proceso de Markov bivariado de dos fases y con coeficientes de deriva no lineales.

## 2.2 Regularidad de un proceso de difusión cambiante

Se sigue del **Teorema 2.1.6** que si los coeficientes satisfacen las condiciones de crecimiento lineal (véase **Definición 1.4.10**) y la condición local de Lipschitz (véase **Definición 1.4.11**), entonces la solución  $(X(t), Y(t))$  a (28) está definida para todo  $t > 0$ . Ahora, supóngase que los coeficientes de deriva y de difusión de  $(X(t), Y(t))$  solo satisfacen la condición local de Lipschitz (39). Definimos los *tiempos de salida*:

**Definición 2.2.1** Para todo  $n = 1, 2, \dots$  y  $(x, \alpha) \in \mathbb{S} \times S$ , definimos  $\beta_n$  como el primer tiempo de salida del proceso  $(X(t), Y(t))$  de la bola abierta centrada en el origen de radio  $n$ ,  $B(0, n) := \{\epsilon \in \mathbb{S} : |\epsilon| \leq n\}$ . Eso es:

$$\beta_n = \beta_n^{x, \alpha} := \inf\{t \geq 0 : |X^{x, \alpha}(t)| \geq n\}, \quad (44)$$

donde  $X^{x, \alpha}(t)$  representa al componente continuo  $X(t)$  del proceso  $(X(t), Y(t))$  que inicia en  $(x, \alpha)$

**Nótese** que la sucesión  $\{\beta_n\}$  es monótona y creciente, por lo que posee un límite (finito o infinito), al cual denotaremos por  $\beta_\infty$ . Además, con el objeto de analizar el comportamiento del proceso dentro (o fuera) de la bola  $B(0, n)$  para  $n$  arbitrariamente grande, vamos a suponer que  $\mathbb{S} = \mathbb{R}$ .

**Definición 2.2.2:** Un proceso de Markov  $(X(t), Y(t))$  que comienza en  $(X(0), Y(0)) = (x, \alpha)$  es regular si:

$$\beta_\infty = \lim_{n \rightarrow \infty} \beta_n = \infty \text{ c.s.} \quad (45)$$

De esta forma también diremos que el proceso  $(X^{x, \alpha}(t), Y^{x, \alpha}(t))$  es regular si y solo si para todo  $T \in (0, \infty)$ , se tiene que:

$$\mathbb{P} \left[ \sup_{0 \leq t \leq T} |X^{x, \alpha}(t)| = \infty \right] = 0. \quad (46)$$

En otras palabras, el proceso es regular si no explota en tiempo finito.

En lo que sigue, analizaremos la regularidad del proceso cuando los coeficientes de deriva y difusión no satisfacen la condición de crecimiento lineal (véase **Definición 1.4.10**). El siguiente resultado otorga condiciones suficientes de regularidad, basándonos en el crecimiento lineal local (véase **Definición 1.4.10**) y la condición de continuidad Lipschitz (véase **Definición 1.4.11**).

**Teorema 2.2.3:** Supóngase que para todo  $i \in S$ , tanto el coeficiente de deriva  $b(\cdot, i)$  como el de difusión  $\sigma(\cdot, i)$  satisfacen las condiciones de crecimiento lineal (véase **Definición 1.4.10**) y la condición de Lipschitz en cualquier conjunto abierto y acotado de  $\mathbb{R}$ , y que también existe una función no negativa  $V(\cdot, \cdot) : \mathbb{R} \times S \rightarrow \mathbb{R}^+$  doblemente diferenciable con respecto a  $x \in \mathbb{R}$  para todo  $i \in S$  tal que existe una  $\gamma_0 > 0$  que satisface lo siguiente:

$$\begin{aligned} \mathcal{L}V(x, i) &\leq \gamma_0 \cdot V(x, i), \text{ para todo } (x, i) \in \mathbb{R} \times S, \\ V_R &:= \inf_{|x| \geq R, i \in S} [V(x, i)], \quad \lim_{R \rightarrow \infty} V_R = \infty. \end{aligned} \quad (47)$$

Entonces el proceso  $(X(t), Y(t))$  es regular.

**Demostración:**

Basta demostrar que (45) se satisface. Supóngase que (45) no se satisface, entonces existirían  $T > 0$  y  $\epsilon > 0$  tales que  $\mathbb{P}_{(x, i)}[\beta_\infty \leq T] > \epsilon$ . Por lo tanto podremos encontrar  $n_1 \in \mathbb{N}$  tal que:

$$\mathbb{P}_{(x, i)}[\beta_n \leq T] > \epsilon, \text{ para todo } n \geq n_1. \quad (48)$$

Defínase:

$$S(x, i, t) = V(x, i)e^{-\gamma_0 t}, \quad (x, i) \in \mathbb{R} \times S, \text{ y } t \geq 0.$$

Entonces  $\left[\frac{\partial}{\partial t} + \mathcal{L}\right] S(x, i, t) \leq 0$ . Gracias a la fórmula de Dynkin (véase **Teorema 2.1.4**) tendremos que:

$$\begin{aligned} & \mathbb{E}_{(x,i)} \left[ V(X(\beta_n \wedge T), Y(\beta_n \wedge T))e^{-\gamma_0(\beta_n \wedge T)} \right] - V(x, i) \\ &= \mathbb{E}_{(x,i)} \left[ \int_0^{\beta_n \wedge T} \left[ \frac{\partial}{\partial t} + \mathcal{L} \right] S(X(u), Y(u), u) du \right] \leq 0. \end{aligned}$$

Por lo que:

$$V(x, i) \geq \mathbb{E}_{(x,i)} \left[ V(X(\beta_n \wedge T), Y(\beta_n \wedge T))e^{-\gamma_0(\beta_n \wedge T)} \right].$$

Nótese que  $\beta_n \wedge T \leq T$  y que  $V$  es no negativa. Por lo tanto:

$$\begin{aligned} V(x, i)e^{\gamma_0 T} &\geq \mathbb{E}_{(x,i)} \left[ V(X(\beta_n \wedge T), Y(\beta_n \wedge T)) \right] \\ &\geq \mathbb{E}_{(x,i)} \left[ V(X(\beta_n), Y(\beta_n)) \mathbb{1}_{\beta_n \leq T} \right]. \end{aligned}$$

Además, ocupando la definición de  $\beta_n$  y (47), tendremos que:

$$V(x, i)e^{\gamma_0 T} \geq V_n \mathbb{P}_{(x,i)} [\beta_n \leq T] > \epsilon V_n \rightarrow \infty \iff n \rightarrow \infty,$$

lo cual es un absurdo. Por lo tanto  $\lim_{n \rightarrow \infty} \beta_n = \infty$  c.s. .

Q.E.D.

Similar al resultado anterior, el siguiente teorema nos otorga condiciones suficientes que debe satisfacer el proceso  $(X(t), Y(t))$  para no ser regular, basándonos en crecimiento lineal local (véase **Definición 1.4.10**) y la condición de continuidad Lipschitz (véase **Definición 1.4.11**).

**Teorema 2.2.4:** *Supóngase que para todo  $i \in S$  tanto el coeficiente de deriva  $b(\cdot, i)$  como el coeficiente de difusión  $\sigma(\cdot, i)$  satisfacen las condiciones de crecimiento lineal y la condición de Lipschitz en cualquier conjunto abierto y acotado de  $\mathbb{R}$ . Supóngase también que existe una función no negativa y acotada  $V(\cdot, \cdot) : \mathbb{R} \times S \rightarrow \mathbb{R}^+$  la cual no es idénticamente cero, para todo  $i \in S$ ,  $V(x, i)$  es doblemente diferenciable con respecto a  $x \in \mathbb{R}$  y existe  $\gamma_1 > 0$  tal que:*

$$\mathcal{L}V(x, i) \geq \gamma_1 V(x, i), \quad \text{para todo } (x, i) \in \mathbb{R} \times S.$$

Entonces el proceso  $(X(t), Y(t))$  no es regular. En particular, para todo  $\epsilon > 0$  tenemos que:

$$\mathbb{P} [\beta_\infty < \kappa + \epsilon | (X(0), Y(0)) = (x_0, l)] > 0,$$

donde  $(x_0, l) \in \mathbb{R} \times S$  satisfacen que  $V(x_0, l) > 0$ , y:

$$\kappa = \frac{1}{\gamma_1} \log \left( \frac{\sup_{(x,i) \in \mathbb{R} \times S} V(x, i)}{V(x_0, l)} \right).$$

(La demostración de este resultado puede encontrarse en [36]).

**Nótese** que si los coeficientes (28) satisfacen las condiciones de crecimiento lineal (38) y la condición local de Lipschitz (39) entonces, (como consecuencia del **Teorema 2.1.6**), existe una solución única  $(X^{x,\alpha}(t), Y^{x,\alpha}(t))$  a (28) para todo  $t \geq 0$ . Por ende, el proceso de difusión cambiante  $(X^{x,\alpha}(t), Y^{x,\alpha}(t))$  es regular. Alternativamente, podemos ocupar el **Teorema 2.2.3** para verificar lo anteriormente mencionado. De hecho, la función:

$$V(x, i) = (|x|^2 + 1)^{\frac{\epsilon}{2}} = (x^2 + 1)^{\frac{\epsilon}{2}}, \quad (x, i) \in \mathbb{R} \times S,$$

satisface todas las condiciones del **Teorema 2.2.3**. Por lo cual se sigue la nota anterior.

En particular, puede verificarse que la función  $V(\cdot, i)$  es doblemente diferenciable para todo  $i \in S$  y que  $\mathcal{L}V(x, i) \leq \gamma_0 V(x, i)$  para algún  $\gamma_0 > 0$  y para todo  $(x, i) \in \mathbb{R} \times S$ . Por lo tanto, para todo  $t > 0$  tendremos que:

$$\mathbb{P}[\beta_n < t | (X(0), Y(0)) = (x_0, i)] \leq \frac{V(x, i)e^{\gamma_0 t}}{V_n},$$

donde  $V_n = (1 + n^2)^{\frac{\epsilon}{2}}$ . Por ende, para todo  $t > 0$  y  $\epsilon > 0$ , existe  $N \in \mathbb{N}$  tal que:

$$\mathbb{P}[\beta_n < t | (X(0), Y(0)) = (x_0, i)] \leq \epsilon, \text{ para todo } n \geq N, \quad (49)$$

uniformemente para todo  $x$  contenido en cualquier conjunto compacto  $F \subset \mathbb{R}$  e  $i \in S$ .

Para finalizar esta sección se presentará una serie de simulaciones programadas en R para evidenciar algunos patrones de comportamiento de los procesos de difusión cambiante, en este caso el análisis de regularidad.

**Ejemplo 2.2.5:** *Considérese el proceso  $(X(t), Y(t))$  que satisface la siguiente ecuación:*

$$dX(t) = X(t) [b(Y(t)) - a(Y(t))X^2(t)dt + \sigma(Y(t))dw(t)], \quad (50)$$

donde  $w(\cdot)$  es un movimiento Browniano unidimensional estándar e  $Y(\cdot) \in S = \{1, 2, \dots, n\}$  es un proceso de salto con generador  $Q(x)$ . Claramente, los coeficientes en (50) no satisfacen la condición de crecimiento lineal a menos que  $a(i) = 0$  para todo  $i \in \{1, 2, \dots, n\}$ . Afirmaremos que si  $a(i) > 0$  para todo  $i \in \{1, 2, \dots, n\}$ , entonces (50) es un proceso regular. Para demostrarlo, vamos a ocupar el **Teorema 2.2.3** y considérese

$$V(x, i) := x^2, \quad (x, i) \in \mathbb{R} \times S.$$

Claramente, la función  $V$  es no negativa y satisface que

$$\lim_{|x| \rightarrow \infty} V(x, i) = \infty, \text{ para todo } i \in S.$$

Por lo que, para aplicar el **Teorema 2.2.3** basta con verificar que para todo  $(x, i) \in \mathbb{R} \times S$ , existe una constante positiva  $K$  tal que  $\mathcal{L}V(x, i) \leq KV(x, i)$ . Para lo cual, se tiene que

$$\mathcal{L}V(x, i) = 2x \cdot x^2 [b(i) - a(i)x^2] + \frac{1}{2} \cdot 2 \cdot x^2 \sigma^2(i) = (2b(i) + \sigma^2(i))x^2 - 2a(i)x^4 \leq KV(x, i),$$

donde  $K := \max\{2b(i) + \sigma^2(i), i \in S\}$ . Nótese que en la ecuación anterior se ocupó que  $a(i) > 0$  para todo  $i \in \{1, 2, \dots, n\}$ . Por lo que se sigue del **Teorema 2.2.3** que el proceso (50) es regular.

Para ilustrar lo anteriormente demostrado, graficaremos una trayectoria de prueba de (50) con los siguientes coeficientes. El proceso de salto  $Y(\cdot) \in S = \{1, 2\}$  es generado por

$$Q(x) = \begin{bmatrix} -3 - \sin(x) \cos(x) & 3 + \sin(x) \cos(x) \\ 2 & -2 \end{bmatrix},$$

además

$$\begin{aligned} b(1) &= 3, a(1) = 2, \sigma(1) = 1 \\ b(2) &= 10, a(2) = 1, \sigma(2) = -1, \end{aligned}$$

y con condiciones iniciales  $(x, i) = (3, 1)$ .

### Ejemplo de Regularidad

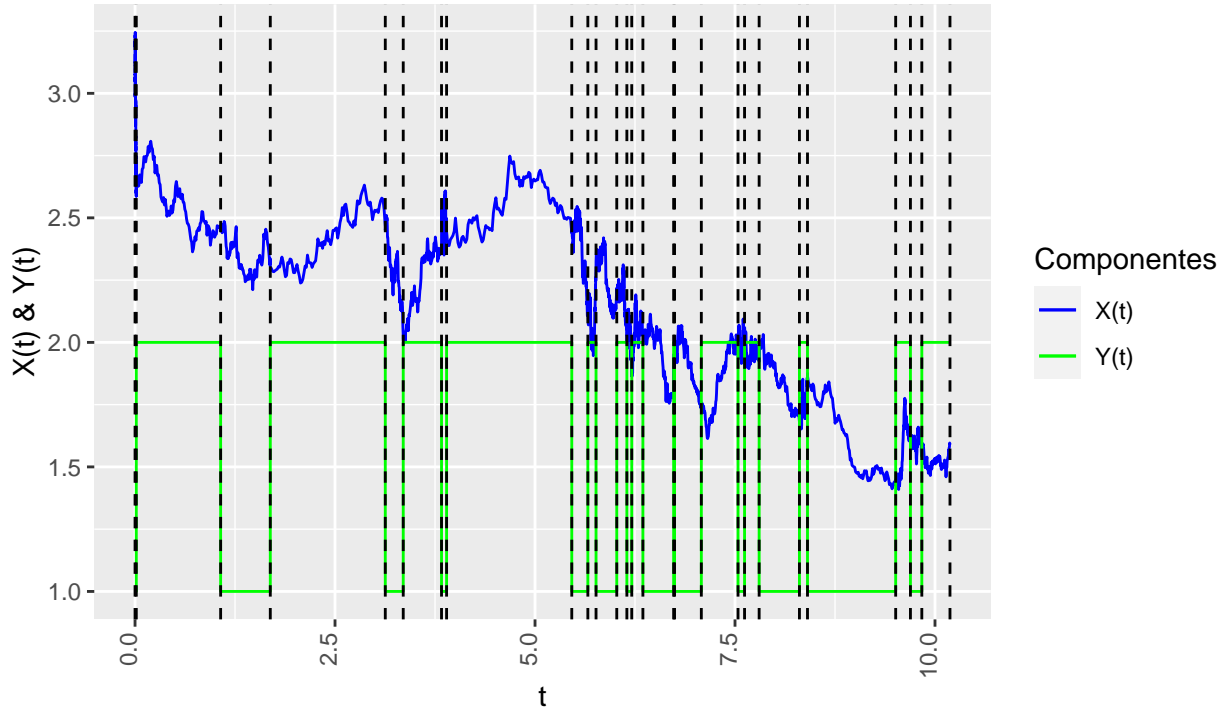


Figura 4: Proceso de Markov bivarido, dos fases y condiciones iniciales  $(x,y)=(3,1)$ .

### 2.3 Continuidad de un proceso de difusión cambiante

Considérese la **Definición 1.4.2**.

Procederemos enunciando una serie de resultados con respecto a la continuidad del proceso  $(X(t), Y(t))$ .

**Teorema 2.3.1:** *Supóngase que se satisfacen las condiciones del Teorema 2.1.6, con la modificación  $Q(x) = Q$  que genera la cadena de Markov independiente al movimiento Browniano. Entonces, el proceso  $Z(t) = (X(t), Y(t))$  es continuo en probabilidad y continuo en media cuadrática.*

**Demostración:**

Vamos a probar que para todo  $\eta > 0$ ,

$$\begin{aligned} \mathbb{P}[|Z(t + \Delta) - Z(t)| \geq \eta] &\rightarrow 0, \text{ si } \Delta \rightarrow 0, \text{ y} \\ \mathbb{E}[|Z(t + \Delta) - Z(t)|^2] &\rightarrow 0, \text{ si } \Delta \rightarrow 0. \end{aligned} \tag{51}$$



Procederemos primero estableciendo la convergencia en media cuadrática. Nótese que:

$$\begin{aligned} Z(t + \Delta) - Z(t) &= (X(t + \Delta), Y(t + \Delta)) - (X(t + \Delta), Y(t)) \\ &\quad + (X(t + \Delta), Y(t)) - (X(t), Y(t)). \end{aligned} \quad (52)$$

El resto de la prueba se divide en varios pasos.

*Paso 1*

Ocupando (52) tenemos que:

$$\mathbb{E} [|Z(t + \Delta) - Z(t)|^2] \leq [\mathbb{E} [|Y(t + \Delta) - Y(t)|^2] + \mathbb{E} [|X(t + \Delta) - X(t)|^2]]. \quad (53)$$

Por lo que para estimar la diferencia del segundo momento basta con considerar dos estimaciones marginales por separado. Lo cual logramos en los siguientes dos pasos.

*Paso 2*

Aseguramos que para cualesquiera  $t \geq 0$  y  $\Delta \geq 0$ ,

$$\mathbb{E} [|X(t + \Delta) - X(t)|^2] \leq K\Delta. \quad (54)$$

Esta estimación, de hecho, es una modificación de los estimadores estándar para ecuaciones diferenciales estocásticas. Ocupa principalmente las condiciones de crecimiento lineal (véase **Definición 1.4.10**) y de Lipschitz (véase **Definición 1.4.11**) de los coeficientes de deriva y difusión y la **Proposición 2.1.7**.

*Paso 3*

Nótese que para todo  $t \geq 0$ ,

$$Y(t) = \sum_{i=1}^N i \mathbb{1}_{\{Y(t)=i\}} = \chi(t) \cdot (1, \dots, N)^*,$$

donde

$$\chi(t) = (\chi_1(t), \dots, \chi_N(t)) = (\mathbb{1}_{\{Y(t)=1\}}, \dots, \mathbb{1}_{\{Y(t)=N\}}) \in \mathbb{R}^{1 \times N}, \quad (55)$$

y  $(1, \dots, N)^* \in \mathbb{R}^N$  es un vector columna. Dado que la cadena de Markov  $Y(t)$  es independiente del movimiento Browniano  $w(\cdot)$  ( $Q$  es una matriz constante) entonces:

$$\chi(t + \Delta) - \chi(t) - \int_t^{t+\Delta} \chi(s) Q ds,$$

es una martingala (véase [6] para mayor referencia). Se sigue que

$$\mathbb{E}_t \left[ \chi(t + \Delta) - \chi(t) - \int_t^{t+\Delta} \chi(s) Q ds \right] = 0,$$

donde  $\mathbb{E}_t$  denota la esperanza condicionada a la  $\sigma$ -álgebra

$$\mathcal{F}_t = \{(X(u), Y(u)) : u \leq t\}.$$

Por lo tanto se sigue que:

$$\left| \int_t^{t+\Delta} \chi(s) Q ds \right| = O(\Delta) \text{ c.s. ,} \quad (56)$$

donde  $O(\Delta)$  satisface que  $\lim_{\Delta \rightarrow \infty} \frac{O(\Delta)}{\Delta} < \infty$ .

Así, obtendremos que:

$$\mathbb{E}_t[\chi(t + \Delta)] = \chi(t) + O(\Delta) \text{ c.s.} \quad (57)$$

Ocupando esta estructura, nótese que:

$$\chi(t + \Delta) - \chi(t) = (\chi_1(t + \Delta) - \chi_1(t), \dots, \chi_N(t + \Delta) - \chi_N(t)),$$

donde  $\chi_i(\cdot)$  está dado por (55). Todo esto junto con (57) implica que:

$$\begin{aligned} \mathbb{E}_t [|\chi_i(t + \Delta) - \chi_i(t)|^2] &= \mathbb{E}_t \left[ \left[ \mathbb{1}_{\{Y(t+\Delta)=i\}} - \mathbb{1}_{\{Y(t)=i\}} \right]^2 \right] \\ &= \left[ \mathbb{E}_t \left[ \mathbb{1}_{\{Y(t+\Delta)=i\}} \right] - 2\mathbb{1}_{\{Y(t)=i\}} \mathbb{E}_t \left[ \mathbb{1}_{\{Y(t+\Delta)=i\}} + \mathbb{1}_{\{Y(t)=i\}} \right] \right] = O(\Delta) \text{ c.s.} \end{aligned} \quad (58)$$

*Paso 4*

Considérese que:

$$\begin{aligned} \mathbb{E} [ |Y(t + \Delta) - Y(t)|^2 ] &= \mathbb{E} [ |\chi(t + \Delta) - \chi(t)| \cdot (1, \dots, N)^* ]^2 \leq K \mathbb{E} [ |\chi(t + \Delta) - \chi(t)|^2 ] \\ &\leq K \sum_{i=1}^N \mathbb{E} [ \mathbb{E}_t [ |\chi_i(t + \Delta) - \chi_i(t)|^2 ] ] \leq K \Delta \rightarrow 0 \text{ si } \Delta \rightarrow 0. \end{aligned} \quad (59)$$

Ocupando (58), (54) y (59), tendremos que (53) implica que:

$$\mathbb{E} [ |Z(t + \Delta) - Z(t)|^2 ] \rightarrow 0 \text{ cuando } \Delta \rightarrow 0.$$

La continuidad en media cuadrada ha sido demostrada, por lo tanto, la continuidad en probabilidad se sigue de la desigualdad de Chebyshev (véase **Proposición 1.4.5**).

Q.E.D.

Ahora extendemos el resultado anterior para el caso en que el proceso dependa de  $x$ , (es decir  $Q$  depende de  $x$ ).

**Teorema 2.3.2:** *Supóngase que se satisfacen las condiciones del Teorema 2.1.6. Entonces, el proceso  $(X(t), Y(t))$  es continuo en probabilidad y continuo en media cuadrática.*

**Demostración:**

Considérese la función  $h(x, \alpha) = \mathbb{1}_{\{\alpha=i\}}$  para cada  $i \in S$ . Dado que  $h$  es independiente de  $x$ , entonces es evidente que  $\mathcal{L}h(x, \alpha) = Q(x)h(x, \cdot)(\alpha)$ . Por ende:

$$\mathbb{E}_t \left[ h(X(t + \Delta), Y(t + \Delta)) - h(X(t), Y(t)) - \int_t^{t+\Delta} \mathcal{L}h(X(s), Y(s)) ds \right] = 0.$$

Sin embargo,

$$\begin{aligned}
& \mathbb{E}_t \left[ h(X(t+\Delta), Y(t+\Delta)) - h(X(t), Y(t)) - \int_t^{t+\Delta} \mathcal{L}h(X(s), Y(s)) ds \right] = \\
& = \mathbb{E}_t \left[ h(X(t+\Delta), Y(t+\Delta)) - h(X(t), Y(t)) - \int_t^{t+\Delta} Q(X(s))h(X(s), \cdot)(Y(s)) ds \right] = \\
& = \mathbb{E}_t \left[ \mathbb{1}_{\{Y(t+\Delta)=i\}} - \mathbb{1}_{\{Y(t)=i\}} - \int_t^{t+\Delta} \sum_{j=1}^N q_{ij}(X(s)) \mathbb{1}_{\{Y(s)=i\}} ds \right].
\end{aligned}$$

Dado que  $Q(x)$  está acotada (y ocupando un argumento similar al de (56)) llegaremos a que:

$$\left| \int_t^{t+\Delta} \chi(s) Q(X(s)) ds \right| = O(\Delta) \text{ c.s.} \quad (60)$$

Finalmente, ocupando (60) se siguen los pasos 3 y 4 de la demostración del **Teorema 2.3.1**. Así se sigue el resultado deseado.

Q.E.D.

Ahora presentaremos un conjunto de definiciones y resultados con respecto a *la propiedad de Feller de los procesos de difusión cambiante*. Esta propiedad indica que cualquier perturbación ‘pequeña’ en la información inicial del proceso debe resultar en una perturbación ‘pequeña’ en el siguiente paso del proceso (véase **Definición 1.2.12**).

Vamos a analizar bajo qué condiciones un proceso de difusión cambiante satisface la propiedad de Feller. En este caso, la función a tomar en consideración para realizar el análisis es  $u(t, x, i) = \mathbb{E}[f(X(t), Y(t)) | (X(0), Y(0)) = (x, i)]$ . Tendremos que probar que  $u(t, x, i)$  es continuo con respecto a la información inicial  $(x, i)$  para todo  $t \geq 0$  y que  $\lim_{t \downarrow 0} u(t, x, i) = f(x, i)$  para cualquier función  $f(\cdot, \cdot)$  medible, continua y acotada. Ocupando el hecho de que el proceso es càdlàg (**Proposición 2.1.8**), y que la función  $f$  es continua y acotada podemos garantizar que:

$$\lim_{t \downarrow 0} u(t, x, i) = \mathbb{E}[f(X(0), Y(0)) | (X(0), Y(0)) = (x, i)] = f(x, i).$$

Nótese que  $u(0, x, i)$  es automáticamente continua debido a la continuidad de  $f$ . Dado que  $S = \{1, \dots, N\}$  es un conjunto finito, basta con probar que  $u(t, x, i)$  es continua con respecto a  $x$  para cualquier  $t > 0$ .

**Lema 2.3.3:** *Supóngase que se satisfacen las condiciones del Teorema 2.1.6, con la simple modificación de que la condición de Lipschitz local (véase Definición 1.4.11) sea reemplazada por la condición de Lipschitz global. Además, supóngase que  $Y(\cdot)$  sea generado por  $Q(x) = Q$  y que  $Y(\cdot)$  es independiente al movimiento Browniano  $w(\cdot)$ . Entonces, para todo  $T > 0$  fijo, tendremos que:*

$$\mathbb{E} \left[ \sup_{0 < t < T} |X^{x_1, i}(t) - X^{x_2, i}(t)|^2 \right] \leq C|x_1 - x_2|^2, \quad (61)$$

donde  $C$  es una constante que depende únicamente de  $T$ , la condición de Lipschitz global y la constante de crecimiento local  $K$ .

(La demostración de este resultado puede encontrarse en [36]).

**Nótese** que una consecuencia inmediata de (61) es garantizar el cumplimiento de la propiedad de Feller para el caso de que el generador de  $Y(t)$  sea constante  $Q$ .

Ahora, para atacar el caso donde el generador  $Q(x)$  ya no sea constante, además de las condiciones del **Teorema 2.1.6**, asumiremos que los coeficientes (28) están acotados, son continuamente diferenciables y que satisfacen la condición de Lipschitz global (todo con respecto a  $x$ ). Es decir, asumiremos que:

**(A1.5.1)**

Para cualesquiera  $\alpha, l \in S$ ,  $i, j = 1, \dots, j$ ,  $a_{ij}(\cdot, \alpha) \in \mathcal{C}^2$ ,  $b(\cdot, \alpha) \in \mathcal{C}^1$ ,  $q_{\alpha, l}(\cdot) \in \mathcal{C}^1$  y alguna constante  $K > 0$ , tenemos que:

$$|b(x, i)| + |\sigma(x, i)| + |q_{ij}(x)| \leq K, x \in \mathbb{R}, i, j \in S, \quad (62)$$

y para todo  $\alpha, l \in S$  y  $x_1, x_2 \in \mathbb{R}$ , tenemos que:

$$|\sigma(x_1, \alpha) - \sigma(x_2, \alpha)| + |b(x_1, \alpha) - b(x_2, \alpha)| + |q_{\alpha, l}(x_1) - q_{\alpha, l}(x_2)| \leq K|x_1 - x_2|. \quad (63)$$

Considérese el siguiente proceso auxiliar  $(Z^{x, \alpha}(t), W^{x, \alpha}(t)) \in \mathbb{R} \times S$ :

$$\begin{cases} dZ(t) = b(Z(t), W(t))dt + \sigma(Z(t), W(t))dw(t) : (Z(0), W(0)) = (x, \alpha) : \alpha \in S, \\ \mathbb{P}[W(t + \Delta) = j | W(t) = i, s \leq t] = \Delta + o(\Delta), i \neq j, \end{cases} \quad (64)$$

donde  $w(t)$  es un movimiento Browniano estándar unidimensional independiente a la cadena de Markov  $W(t)$ . Dicha cadena posee el siguiente generador:

$$\begin{bmatrix} -(m_0 - 1) & 1 & \dots & 1 & 1 \\ 1 & -(m_0 - 1) & \dots & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 & -(m_0 - 1) \end{bmatrix} = -m_0 I_{m_0} + (\mathbb{1}_{m_0}, \mathbb{1}_{m_0}, \dots, \mathbb{1}_{m_0}),$$

donde  $I_{m_0}$  es la matriz identidad de dimensión  $m_0 \times m_0$  e  $\mathbb{1}_{m_0}$  es un vector columna de dimensión  $m_0$  con componentes todas son iguales a 1.

Para todo  $T > 0$ , definimos  $\mu_{T,1}(\cdot)$  como la medida inducida por el proceso  $(X^{x, \alpha}(t), Y^{x, \alpha}(t))$  y  $\mu_{T,2}(\cdot)$  como la medida inducida por el proceso auxiliar  $(Z^{x, \alpha}(t), W^{x, \alpha}(t))$  (es decir, estas dos medidas  $\mu_{T,1}(\cdot)$  y  $\mu_{T,2}(\cdot)$  se definen a partir de la función de probabilidad de los procesos  $(X^{x, \alpha}(t), Y^{x, \alpha}(t))$  y  $(Z^{x, \alpha}(t), W^{x, \alpha}(t))$  respectivamente), por lo que  $\mu_{T,1}(\cdot)$  y  $\mu_{T,2}(\cdot)$  son medidas de probabilidad. Entonces (ocupando los resultados obtenidos en [16] y [15])  $\mu_{T,1}(\cdot)$  es absolutamente continua con respecto a  $\mu_{T,2}(\cdot)$ . Además, la correspondiente derivada de Radon-Nikodym tiene la siguiente forma:

$$\begin{aligned} p_T(Z^{x, \alpha}(\cdot), W^{x, \alpha}(\cdot)) &= \frac{d\mu_{T,1}}{d\mu_{T,2}}(Z^{x, \alpha}(\cdot), W^{x, \alpha}(\cdot)) \\ &= e^{(m_0 - 1)T} \exp \left\{ - \int_{\tau_n(T)}^T q_{r(\tau_n(T))}(Z(s)) ds \right\} \times \prod_{i=0}^{n(T)-1} \left[ q_{r(\tau_i)r(\tau_{i+1})}(Z(\tau_{i+1})) \exp \left\{ - \int_{\tau_i}^{\tau_{i+1}} q_{r(\tau_i)}(Z(s)) ds \right\} \right], \end{aligned} \quad (65)$$

donde para todo  $i \in S$ ,  $q_i(x) := -q_{ii}(x) = \sum_{j \neq i} q_{ij}(x)$ ,  $\tau_i$  es una sucesión de tiempos de paro definida por:

$$\tau_0 = 0,$$

y para todo  $i = 0, 1, \dots$ ,

$$\tau_{i+1} = \inf\{t > \tau_i : W(t) \neq W(\tau_i)\},$$

y  $n(T) = \max\{n \in \mathbb{N} : \tau_n \leq T\}$ . Nótese que si  $n(T) = 0$ , entonces:

$$p_T(Z^{x,\alpha}(\cdot), W^{x,\alpha}(\cdot)) = \exp\left\{-\int_{\tau_n(T)}^T [q_\alpha(Z(s)) - m_0 + 1](Z(s))ds\right\}.$$

Con respecto a la derivada de Radon-Nikodyn  $p_T(Z^{x,\alpha}(\cdot), W^{x,\alpha}(\cdot))$ , tenemos los siguientes resultados:

**Lema 2.3.4:** *Supóngase que se satisfacen las condiciones del Teorema 2.1.6 y también (A1.5.1). Entonces, para todo  $T > 0$  y  $(a, \alpha) \in \mathbb{R} \times S$ , tenemos que:*

$$\mathbb{E}[|p_T(Z^{x,\alpha}(\cdot), W^{x,\alpha}(\cdot))|] \leq K < \infty.$$

(La demostración de este resultado puede encontrarse en [36]).

**Lema 2.3.5** *Supóngase que se satisfacen las condiciones del Lema 2.3.4. Entonces, para todo  $T > 0, x_1, x_2 \in \mathbb{R}$  y  $\alpha \in S$ , tenemos que:*

$$\mathbb{E}[|p_T(Z^{x_1,\alpha}(\cdot), W^\alpha(\cdot)) - p_T(Z^{x_2,\alpha}(\cdot), W^\alpha(\cdot))|] \leq K|x_1 - x_2|.$$

(La demostración de este resultado puede encontrarse en [36]).

Con todos los resultados anteriores, ya podemos probar la Propiedad de Feller para los procesos de difusión cambiante.

**Teorema 2.3.6** *Sea  $(X^{x,\alpha}(t), Y^{x,\alpha}(t))$  la solución del sistema dado por (28)-(29), con  $(X(0), Y(0)) = (x, \alpha)$ . Supóngase que se satisfacen las condiciones del Teorema 2.1.6. Entonces para cualquier función continua y acotada  $g(\cdot, \cdot) : \mathbb{R} \times S \rightarrow \mathbb{R}$ , la función  $u(x, \alpha) = \mathbb{E}_{(x,\alpha)}[g(X(t), Y(t))] = \mathbb{E}[g(X^{x,\alpha}(t), Y^{x,\alpha}(t))]$  es continua con respecto a  $x$ .*

**Demostración:**

La prueba de este teorema consiste en dos pasos. El primer paso corresponde al caso donde suponemos que se satisface (A1.5.1). El caso general se trabaja en el segundo paso.

*Paso 1*

Suponemos que se satisface (A1.5.1). Entonces para una  $t > 0$  fija, tendremos que:

$$\mathbb{E}[g(X^{x,\alpha}(t), Y^{x,\alpha}(t))] = \mathbb{E}[g(Z^{x,\alpha}(t), W^\alpha(t)) p_t(Z^{x,\alpha}(\cdot), W^\alpha(\cdot))].$$

Sea  $\{x_n\}$  una sucesión de puntos que converge a  $x$ . Entonces el Lema 2.3.3 implica que:

$$\mathbb{E}[|Z^{x_n,\alpha}(t) - Z^{x,\alpha}(t)|^2] \rightarrow 0 \text{ cuando } n \rightarrow \infty.$$

Por lo que existe una subsucesión  $\{y_n\}$  de  $\{x_n\}$  tal que:

$$Z^{y_n,\alpha}(t) \rightarrow Z^{x,\alpha}(t) \text{ c.s. cuando } n \rightarrow \infty.$$

En virtud del Lema 2.3.5:

$$\mathbb{E}[|p_t(Z^{y_n,\alpha}(\cdot), W^\alpha(\cdot)) - p_t(Z^{x,\alpha}(\cdot), W^\alpha(\cdot))|] \rightarrow 0 \text{ cuando } n \rightarrow \infty.$$

Por lo que existe una subsucesión  $\{z_n\}$  de  $\{y_n\}$  tal que:

$$p_t(Z^{z_n, \alpha}(\cdot), W^\alpha(\cdot)) \rightarrow p_t(Z^{x, \alpha}(\cdot), W^\alpha(\cdot)) \text{ c.s. cuando } n \rightarrow \infty.$$

Dado que  $g$  está acotada y es continua, se sigue del **Lema 2.3.4** y del Teorema de convergencia dominada (véase **Teorema 1.4.15**) que:

$$\begin{aligned} u(x, \alpha) &= \mathbb{E}[g(X^{x, \alpha}(t), Y^{x, \alpha}(t))] = \mathbb{E}[g(Z^{x, \alpha}(t), W^\alpha(t))p_t(Z^{x, \alpha}(\cdot), W^\alpha(\cdot))] \\ &= \mathbb{E}\left[\lim_{n \rightarrow \infty} g(Z^{z_n, \alpha}(t), W^\alpha(t))p_t(Z^{z_n, \alpha}(\cdot), W^\alpha(\cdot))\right] = \lim_{n \rightarrow \infty} \mathbb{E}[g(X^{z_n, \alpha}(t), Y^{z_n, \alpha}(t))] = \lim_{n \rightarrow \infty} u(y_n, \alpha). \end{aligned}$$

Por lo tanto, toda sucesión  $\{x_n\}$  que converge a  $x$  posee una subsucesión  $\{z_n\}$  tal que  $u(x, \alpha) \leq \liminf_{n \rightarrow \infty} u(z_n, \alpha)$ . Puede demostrarse que el conjunto  $\{x \in \mathbb{R} : u(x, \alpha) > \beta\}$  es abierto para todo  $\beta$  real y cualquier  $\alpha \in S$ . Esto es,  $u(\cdot, \alpha)$  es semi-continua por abajo para todo  $\alpha \in S$ .

Aplicando el anterior argumento a  $-u(x, \alpha) = \mathbb{E}[-g(X^{x, \alpha}(t), Y^{x, \alpha}(t))]$ , tendremos que:

$$u(x, \alpha) \geq \limsup_{n \rightarrow \infty} u(z_n, \alpha),$$

por lo que  $u(\cdot, \alpha)$  es semi-continua por arriba para todo  $\alpha \in S$ . Por lo tanto,  $u(x, \alpha)$  es continua con respecto a  $x$ .

*Paso 2*

Fije  $t \geq 0$ . Sea  $N$  cualquier entero positivo, definimos el proceso  $N$ -truncado  $X^N(t)$  de forma que  $X^N(t) = X(t)$  hasta el primer tiempo de salida de la  $N$ -bola  $B(0, N) = \{x \in \mathbb{R} : |x| < N\}$ . Asociado al proceso  $N$ -truncado  $X^N(t)$  construiremos un operador auxiliar como sigue. Para  $N = 1, 2, \dots$ , sea  $\phi^N(x)$  una función de la clase  $C^\infty$  de rango  $[0, 1]$  que satisface:

$$\phi^N(x) = \begin{cases} 1, & \text{si } |x| \leq N, \\ 0, & \text{si } |x| \geq N + 1. \end{cases}$$

Ahora para  $j, k = 1, \dots, n$  y  $(x, i) \in \mathbb{R} \times S$ , definimos:

$$\begin{aligned} a_{jk}^N(x, i) &:= a_{jk}(x, i)\phi^N(x), \\ b_j^N(x, i) &:= b_j(x, i)\phi^N(x). \end{aligned}$$

Para toda  $\psi(\cdot, i) \in C^2, i \in S$ , definimos el operador  $\mathcal{L}^N$  como

$$\begin{aligned} \mathcal{L}^N \psi(x, i) &= \frac{1}{2} \sum_{j, k=1}^n a_{jk}^N(x, i) \frac{\partial^2}{\partial x_j \partial x_k} \psi(x, i) \\ &\quad + \sum_{j=1}^n b_j^N(x, i) \frac{\partial}{\partial x_j} \psi(x, i) + \sum_{j=1}^n q_{ij}(x) \psi(x, j). \end{aligned} \tag{66}$$

Denotaremos por  $\mathbb{P}_{(x, i)}^N$  a la medida de probabilidad asociada al problema de martingalas cuyo operador es  $\mathcal{L}^N$  con coeficientes  $a_{jk}^N(x, i), b_j^N(x, i)$  y  $q_{ij}(x)$ , también denotaremos por  $\mathbb{E}_{(x, i)}^N$  a su correspondiente media. Entonces, ocupando el *Paso 1* podemos concluir que  $(X^N(t), Y(t))$  es Feller.

Tal y como se define en (44), sea  $\beta_N$  el primer tiempo de salida de la bola  $B(0, N)$ . Entonces en virtud del resultado de unicidad fuerte en [34], las probabilidades  $\mathbb{P}_{(x, i)}$  y  $\mathbb{P}_{(x, i)}^N$  coinciden hasta el momento en el que

el componente continuo del proceso alcanza la frontera  $|x| = N$ . Por lo tanto se sigue que para toda función acotada y Borel-medible  $f(\cdot, \cdot) : \mathbb{R} \times S \rightarrow \mathbb{R}$ , tenemos que:

$$\mathbb{E}_{(x,i)} [f(X(t), Y(t)) \mathbb{1}_{\beta_N > t}] = \mathbb{E}_{(x,i)}^N [f(X(t), Y(t)) \mathbb{1}_{\beta_N > t}]. \quad (67)$$

(Alternativamente, uno puede obtener (67) al demostrar que el problema de martingala con operador  $\mathcal{L}$  tiene solución única en el sentido débil, véase Lema 7.18 de [35]).

Fijando  $(x_0, \alpha) \in \mathbb{R} \times S$ , en virtud de (49) y (67), se sigue que para todo  $t > 0$  y  $\epsilon > 0$ , existe  $N \in \mathbb{N}$  suficientemente grande y una vecindad acotada  $N(x_0)$  de  $x_0$  ( $N(x_0) = \{x : |x - x_0| \leq N\}$ ) tal que para todo  $(x, i) \in N(x_0) \times S$ , tenemos que:

$$\begin{aligned} |\mathbb{E}_{(x,i)} [g(X(t), Y(t))] - \mathbb{E}_{(x,i)}^N [g(X(t), Y(t))]| &= |\mathbb{E}_{(x,i)} [g(X(t), Y(t)) \mathbb{1}_{\beta_n \leq t}] - \mathbb{E}_{(x,i)}^N [g(X(t), Y(t)) \mathbb{1}_{\beta_n \leq t}]| \\ &\leq \|g\|_\infty [\mathbb{P}_{(x,i)} [\beta_n \leq t] + \mathbb{P}_{(x,i)}^N [\beta_n \leq t]] = 2\|g\|_\infty [1 - \mathbb{P}_{(x,i)} [\beta_n > t]] < \frac{\epsilon}{2}, \end{aligned} \quad (68)$$

donde  $\|\cdot\|_\infty$  es la norma del supremo.

Ahora sea  $\{x_n\} \subset N(x_0)$  cualquier sucesión que converge a  $x_0$ , entonces se sigue de (68) y del *Paso 1* que para  $N$  suficientemente grande,

$$\begin{aligned} &|\mathbb{E}[g(X^{x_n, \alpha}(t), Y^{x_n, \alpha}(t))] - \mathbb{E}[g(X^{x_0, \alpha}(t), Y^{x_0, \alpha}(t))]| \\ &\leq |\mathbb{E}[g(X^{x_n, \alpha}(t), Y^{x_n, \alpha}(t))] - \mathbb{E}^N[g(X^{x_n, \alpha}(t), Y^{x_n, \alpha}(t))]| \\ &+ |\mathbb{E}^N[g(X^{x_n, \alpha}(t), Y^{x_n, \alpha}(t))] - \mathbb{E}^N[g(X^{x_0, \alpha}(t), Y^{x_0, \alpha}(t))]| \\ &+ |\mathbb{E}^N[g(X^{x_0, \alpha}(t), Y^{x_0, \alpha}(t))] - \mathbb{E}[g(X^{x_0, \alpha}(t), Y^{x_0, \alpha}(t))]| \\ &\leq \epsilon + |\mathbb{E}^N[g(X^{x_n, \alpha}(t), Y^{x_n, \alpha}(t))] - \mathbb{E}[g(X^{x_0, \alpha}(t), Y^{x_0, \alpha}(t))]| \\ &\rightarrow_{n \rightarrow \infty} \epsilon + 0 = \epsilon. \end{aligned}$$

Con lo cual hemos demostrado que si  $n \rightarrow \infty$

$$\begin{aligned} u(x_n, \alpha) &= \mathbb{E}[g(X^{x_n, \alpha}(t), Y^{x_n, \alpha}(t))] \\ &\rightarrow \mathbb{E}[g(X^{x_0, \alpha}(t), Y^{x_0, \alpha}(t))] = u(x_0, \alpha) \end{aligned}$$

tal y como se deseaba.

Q.E.D.

**Corolario 2.3.7** *Supóngase que se satisfacen las condiciones del Teorema 2.3.6. Entonces, el proceso  $(X(t), Y(t))$  satisface la propiedad fuerte de Markov.*

**Demostración:**

Este resultado se sigue directamente del Teorema 2.24 y Lema 2.4 de [17] y del **Teorema 2.3.6**.

Q.E.D.

**Proposición 2.3.8:** *Bajo las condiciones del Teorema 2.2.3, la ecuación (28) junto con (29) tiene una solución única casi seguramente.*

**Demostración:**

En virtud del **Teorema 2.2.3** el proceso  $(X(t), Y(t))$  es regular y

$$\mathbb{P}[\beta_\infty = \infty] = 1. \quad (69)$$

Definimos el proceso  $N$  – truncado  $X^N(t)$  como en el *Paso 2* de la demostración del **Teorema 2.3.6**. En virtud del **Teorema 2.1.6**, la condición de crecimiento lineal local (véase **Definición 1.4.10**) y la condición de Lipschitz (véase **Definición 1.4.11**),  $X^N(t)$  es la única solución de (28)-(29) para todo  $t < \beta_N$ . La regularidad implica que  $X^N(t)$  es una solución de (28)-(29) para todo  $t \geq 0$ . Dado que  $X^N(t)$  coincide con  $X(t)$  para todo  $t < \beta_N$ ,

$$\mathbb{P} \left[ \sup_{0 \leq t \leq \beta_N} |X^N(t) - X(t)| > 0 \right] = 0.$$

Ocupando (69) y haciendo  $N \rightarrow \infty$ , se sigue el resultado deseado.

Q.E.D.

Por otro lado, también nos es posible realizar un análisis con respecto a *la propiedad fuerte de Feller para procesos de difusión cambiante* (véase **Definición 1.2.13**).

Supóngase válidas las siguientes condiciones por el resto de la sección.

**(A1.5.2)**

Para todo  $\alpha$  e  $i \in S$ , los coeficientes  $b_j(x, i)$ ,  $\sigma_{jk}(x, i)$  y  $q_{i\alpha}(x)$  son Hölder continuos con exponente  $\gamma : 0 < \gamma \leq 1$  (véase **Definición 1.4.3**).

**(A1.5.3)**

La función matricial  $Q(x)$  es irreducible para todo  $x \in \mathbb{R}^r$  (es decir, que el proceso  $Y(\cdot)$  cuya matriz de intensidad de transición es  $Q(x)$  satisface que todos los estados  $S = \{1, \dots, N\}$  son alcanzables desde cualquier estado  $i \in S$ ).

**(A1.5.4)**

Para todo  $i \in S$ ,  $a(x, i) = [a_{jk}(x, i)]$  es simétrica y satisface (para alguna constante positiva  $\kappa \in \mathbb{R}$  y para todo  $x \in \mathbb{R}^r$ ):

$$\langle a(x, i)\epsilon, \epsilon \rangle \geq \kappa|\epsilon|^2, \text{ para todo } \epsilon \in \mathbb{R}^r. \quad (70)$$

Recuérdese que  $r = 1$  para el caso  $\mathbb{S} = (a, b) \subset \mathbb{R}$ , por lo que para este caso se reescribe como:

$$(a(x, i) \cdot \epsilon) \cdot (\epsilon) \geq \kappa|\epsilon|^2, \text{ para todo } \epsilon \in \mathbb{R}, \quad (71)$$

de manera que  $a(x, i) \in \mathbb{R}$ , de forma que la condición de simetría de  $a(x, i)$  es innecesaria.



**Lema 2.3.9:** Supóngase que  $f(\cdot, i) \in C_b(\mathbb{R})$  (la clase de funciones continuas y acotadas sobre  $\mathbb{R}$ ) tal que  $i \in S$ . Si para todo  $i \in S$  la función  $u(\cdot, \cdot, i) \in C^{1,2}(\mathbb{R} \times \mathbb{R})$  (la clase de funciones cuyas derivadas con respecto a  $t$  son continuas y sus derivadas parciales con respecto a  $x$  hasta el segundo orden son continuas) es acotada y satisface:

$$\begin{cases} \frac{\partial u}{\partial t} = \mathcal{L}u, & t > 0, (x, i) \in \mathbb{R} \times S. \\ u(0, x, i) = f(x, i), & (x, i) \in \mathbb{R} \times S. \end{cases} \quad (72)$$

Entonces  $u(x, t, i) = \mathbb{E}_{(x,i)} [f(X(t), Y(t))]$ .

(La demostración de este resultado puede encontrarse en [36]).

Considérese la **Definición 1.2.13** sobre la propiedad fuerte de Feller.

**Lema 2.3.10** Supóngase que, además de satisfacerse las condiciones **(A1.5.2)**-**(A1.5.4)**, se satisface que para todo  $i, l \in S$  y  $j, k = 1$  los coeficientes  $a_{jk}(x, i), b_j(x, i)$  y  $q_{il}(x)$  están acotados. Entonces, el proceso  $(X(t), Y(t))$  satisface la propiedad fuerte de Feller.

(La demostración de este resultado puede encontrarse en [36]).

**Lema 2.3.11:** Supóngase que el proceso  $(X(t), Y(t))$  satisface propiedad fuerte de Feller. Denotaremos  $U := D \times J \subset \mathbb{R} \times S$ , donde  $D$  es un subconjunto no vacío abierto de  $\mathbb{R}$ . Entonces para todo  $t > 0$  y para toda función  $f(\cdot, \cdot)$  acotada, real Borel medible en  $U$  las funciones

$$\begin{aligned} F(x, i) &:= \mathbb{E}_{(x,i)} [\mathbb{1}_{\{\tau_u > t\}} f(X(t), Y(t))], \\ G(x, i) &:= \mathbb{E}_{(x,i)} [f(X(t \wedge \tau_u), Y(t \wedge \tau_u))], \end{aligned}$$

son continuas en  $U$ , donde  $\tau_u$  es el primer tiempo de salida de  $U$ .

(La demostración de este resultado puede encontrarse en [36]).

Con todos los resultados previos, presentamos el resultado principal de esta sección:

**Teorema 2.3.12** Supóngase que se satisfacen las condiciones **(A1.5.2)**-**(A1.5.4)**. Entonces el proceso  $(X(t), Y(t))$  satisface la propiedad fuerte de Feller.

**Demostración:**

Tal como en la prueba del **Teorema 2.3.6**, denotaremos por  $\mathbb{P}_{(x,i)}^N$  a la medida de probabilidad para la cual el problema de martingalas [véase 32] tiene un operador  $\mathcal{L}^N$  definido en (66) y denotaremos por  $\mathbb{E}_{(x,i)}^N$  a su correspondiente media. Entonces el **Lema 2.3.10** implica que  $\mathbb{P}_{(x,i)}^N$  satisface la condición fuerte de Feller. Tal como en (44), sea  $\beta_N$  el primer tiempo de salida de la bola  $B(0, N) = \{x \in \mathbb{R} : |x| < N\}$ . Entonces, tal y como se argumenta en la prueba del **Teorema 2.3.6**, para cualquier función acotada y Borel medible  $f(\cdot, \cdot) \in \mathbb{R} \times S$  y  $(x, i) \in \mathbb{R} \times S$  tenemos que:

$$\mathbb{E}_{(x,i)} [f(X(t), Y(t)) \mathbb{1}_{\beta_n > t}] = \mathbb{E}_{(x,i)}^N [f(X(t), Y(t)) \mathbb{1}_{\beta_n > t}].$$

Se sigue que:

$$\begin{aligned} \left| \mathbb{E}_{(x,i)} [f(X(t), Y(t)) \mathbb{1}_{\beta_n > t}] - \mathbb{E}_{(x,i)}^N [f(X(t), Y(t)) \mathbb{1}_{\beta_n > t}] \right| &= \left| \mathbb{E}_{(x,i)} [f(X(t), Y(t)) \mathbb{1}_{\beta_n \leq t}] - \mathbb{E}_{(x,i)}^N [f(X(t), Y(t)) \mathbb{1}_{\beta_n \leq t}] \right| \\ &= \left| \int_{\Omega} f(X(t), Y(t)) \mathbb{1}_{\beta_n \leq t} [\mathbb{P}_{(x,i)}(dw) - \mathbb{P}_{(x,i)}^N(dw)] \right| \leq 2 \|f\|_{\infty} [1 - \mathbb{P}_{(x,i)}^N[\beta_n > t]], \end{aligned} \quad (73)$$

donde  $\|\cdot\|_\infty$  es la norma del supremo. Fijando  $(x_0, i) \in \mathbb{R} \times S$ , debido a que el proceso  $(X(t), Y(t))$  es regular, por (45),  $\beta_N \rightarrow \infty$  c.s.  $\mathbb{P}_{(x_0, i)}$  cuando  $N \rightarrow \infty$  (es decir,  $\beta_N \rightarrow \infty$  casi seguramente en probabilidad cuando  $X(0) = x_0$  e  $Y(0) = i$ ). Por lo tanto para cualquier  $\epsilon > 0$ , podemos escoger algún  $N$  suficientemente grande de forma que:

$$1 - \mathbb{P}_{(x_0, i)}^N[\beta_N > t] = 1 - \mathbb{P}_{(x_0, i)}[\beta_N > t] < \frac{\epsilon}{12\|f\|_\infty}. \quad (74)$$

Además, ocupando el **Lema 2.3.11**, la función  $x \mapsto \mathbb{P}_{(x, i)}^N[\beta_N > t]$  es continua. Por ende existe  $\delta_1 > 0$  tal que si  $|x - x_0| < \delta_1$  entonces:

$$1 - \mathbb{P}_{(x_0, i)}^N[\beta_N > t] \leq 1 - \mathbb{P}_{(x, i)}[\beta_N > t] + |\mathbb{P}_{(x_0, i)}[\beta_N > t] - \mathbb{P}_{(x, i)}[\beta_N > t]| < \frac{\epsilon}{6\|f\|}. \quad (75)$$

Nótese que también en virtud del **Lema 2.3.10**, existe una  $\delta_2 > 0$  tal que:

$$\left| \mathbb{E}_{(x, i)}^N[f(X(t), Y(t))] - \mathbb{E}_{(x_0, i)}^N[f(X(t), Y(t))] \right| < \frac{\epsilon}{3} \text{ cuando } |x - x_0| < \delta_2. \quad (76)$$

Finalmente, se sigue de (73)-(76) que si  $|x - x_0| < \delta$ , donde  $\delta = \min\{\delta_1, \delta_2\}$ , tendremos que:

$$\begin{aligned} & \left| \mathbb{E}_{(x, i)}[f(X(t), Y(t))] - \mathbb{E}_{(x_0, i)}[f(X(t), Y(t))] \right| \\ & \leq \left| \mathbb{E}_{(x, i)}[f(X(t), Y(t))] - \mathbb{E}_{(x, i)}^N[f(X(t), Y(t))] \right| \\ & \quad + \left| \mathbb{E}_{(x_0, i)}[f(X(t), Y(t))] - \mathbb{E}_{(x_0, i)}^N[f(X(t), Y(t))] \right| \\ & \quad + \left| \mathbb{E}_{(x, i)}^N[f(X(t), Y(t))] - \mathbb{E}_{(x_0, i)}^N[f(X(t), Y(t))] \right| < \epsilon. \end{aligned}$$

Con lo cual se concluye la prueba del teorema.

Q.E.D.

**Teorema 2.3.13** *Supóngase que se satisfacen las condiciones (A1.5.2)-(A1.5.4). Sea  $U := D \times J \subset \mathbb{R} \times S$ , donde  $D$  es un subconjunto no vacío abierto de  $\mathbb{R}$  con una frontera lo suficientemente suave. Entonces el proceso  $(X(\cdot \wedge \tau_u), Y(\cdot \wedge \tau_u))$  satisface la propiedad fuerte de Feller, donde  $\tau_u$  es el primer tiempo de salida de  $U$ .*

**Demostración:**

Una ligera modificación al argumento estándar en el Teorema 13.5 de [14] es suficiente para demostrar este resultado.

Q.E.D.

**Definición 2.3.14** *Sea  $U := D \times J \subset \mathbb{R} \times S$ , donde  $D$  es un subconjunto no vacío de  $\mathbb{R}$ . Una función Borel medible  $u : U \rightarrow \mathbb{R}$  es  $\mathcal{L}$ -armónica en  $U$  si  $u$  está acotada en subconjuntos compactos de  $U$  y si para todo  $(x, i) \in U$  y cualquier  $V = D_1 \times J_2$  donde  $D_1 \subset D$  es una vecindad de  $x$  e  $i \in J_2 \subset J$  tenemos que:*

$$u(x, i) = \mathbb{E}_{(x, i)}[u(X(\tau_V), Y(\tau_V))],$$

donde  $\tau_V$  representa el primer tiempo de salida del proceso  $(X(t), Y(t))$  de  $V$ .

**Nótese** que si el proceso  $(X(t), Y(t))$  satisface la propiedad fuerte de Feller, entonces cualquier función  $\mathcal{L}$ -armónica con respecto a  $(X(t), Y(t))$  es continua.

**Proposición 2.3.15** *Supóngase que se satisfacen las condiciones (A1.5.2)-(A1.5.4). Entonces cualquier función  $\mathcal{L}$ -armónica con respecto a  $(X(t), Y(t))$  es continua en  $U$ .*

*(La demostración de este resultado puede encontrarse en [36]).*

Un problema para el proceso de difusión cambiante está bien definido si existe una solución única al problema de valores iniciales y la solución depende de forma continua de las condiciones iniciales. La dependencia continua en la información inicial puede obtenerse ocupando la propiedad de Feller al escoger la función  $u(\cdot)$  adecuadamente. En el subsecuente desarrollo también ocuparemos la dependencia suave en la información inicial, lo cual es más difícil de obtener.

Un vector  $\beta = (\beta_1, \dots, \beta_n)$  de entradas enteras no negativas es un *multi-índice*, donde

$$|\beta| = \beta_1 + \dots + \beta_n,$$

y definimos  $D_x^\beta$  como:

$$D_x^\beta = \frac{\partial^\beta}{\partial x^\beta} = \frac{\partial^\beta}{\partial x_1^{\beta_1} \dots \partial x_r^{\beta_r}}.$$

Se tiene entonces el siguiente teorema:

**Teorema 2.3.16** *Sea  $(X^{x,\alpha}(t), Y^{x,\alpha}(t))$  la solución al sistema dado por (28) y (29). Supóngase que las condiciones del Teorema 2.1.6 se satisfacen para todo  $i \in S$ ,  $b(\cdot, i)$  y  $\sigma(\cdot, i)$  tienen derivadas parciales continuas con respecto a la variable  $x$  hasta el segundo orden y que:*

$$\left| D_x^\beta b(x, i) \right| + \left| D_x^\beta \sigma(x, i) \right| \leq K(1 - |x|^\gamma), \quad (77)$$

donde  $K$  y  $\gamma$  son constantes positivas y  $\beta$  es un multi-índice con  $|\beta| \leq 2$ . Entonces  $X^{x,\alpha}(t)$  es continuamente diferenciable en media cuadrada con respecto a  $x$ .

Los siguientes resultados serán la base de la demostración del **Teorema 2.3.16**.

**Definición 2.3.17** *Sea  $\Delta \neq 0$ , denotemos por  $\tilde{x} = x + \Delta$ . Tal y como se trabaja en el Lema 2.3.3, sea  $(X(t), Y(t))$  el proceso de difusión cambiante que satisface (28) y (29) con condiciones iniciales  $(x, \alpha)$  y  $(\tilde{X}(t), \tilde{Y}(t))$  el proceso que comienza en  $(\tilde{x}, \alpha)$ , es decir  $(X(0), Y(0)) = (x, \alpha)$  y  $(\tilde{X}(0), \tilde{Y}(0)) = (\tilde{x}, \alpha)$ .*

*Fijamos  $T > 0$  y sea  $0 < t < T$ . Definimos:*

$$Z^\Delta(t) = Z^{x,\Delta,\alpha}(t) := \frac{\tilde{X}(t) - X(t)}{\Delta}. \quad (78)$$

*Por lo tanto:*

$$\begin{aligned} Z^\Delta(t) &= 1 + \phi^\Delta(t) + \frac{1}{\Delta} \int_0^t [b(\tilde{X}(s), Y(s)) - b(X(s), Y(s))] ds \\ &\quad + \frac{1}{\Delta} \int_0^t [\sigma(\tilde{X}(s), Y(s)) - \sigma(X(s), Y(s))] dw(s), \end{aligned} \quad (79)$$

donde

$$\phi^\Delta(t) = \frac{1}{\Delta} \int_0^t [b(\tilde{X}(s), \tilde{Y}(s)) - b(\tilde{X}(s), Y(s))] ds + \frac{1}{\Delta} \int_0^t [\sigma(\tilde{X}(s), \tilde{Y}(s)) - \sigma(\tilde{X}(s), Y(s))] dw(s). \quad (80)$$

**Lema 2.3.18:** *Bajo las condiciones del Teorema 2.3.16,*

$$\lim_{\Delta \rightarrow 0} \mathbb{E} \left[ \sup_{t \in [0, T]} \left| \phi^\Delta(t) \right|^2 \right] = 0.$$

(La demostración de este resultado puede encontrarse en [36]).

**Proposición 2.3.19:** *Bajo las condiciones del Lema 2.3.3, exceptuando que a  $Q(x)$  se le permita ser  $x$ -dependiente, la conclusión del Lema 2.3.3 sigue siendo válida.*

(La demostración de este resultado puede encontrarse en [36]).

Se sigue de inmediato el siguiente resultado:

**Corolario 2.3.20:** *Bajo las condiciones del Teorema 2.3.16,  $X^{x, \alpha}(t)$  es continua en media cuadrada con respecto a  $x$ .*

**Demostración del Teorema 2.3.16:**

Ocupando el **Lema 2.3.18** y **Proposición 2.3.19**, procederemos a probar el **Teorema 2.3.16**. Dado que  $b(\cdot, j)$  es doblemente diferenciable y continua con respecto a  $x$ , podemos escribir:

$$\begin{aligned} & \frac{1}{\Delta} \int_0^t \left[ b(\tilde{X}(s), \tilde{Y}(s)) - b(\tilde{X}(s), Y(s)) \right] ds \\ &= \frac{1}{\Delta} \int_0^t \int_0^1 \left[ \frac{d}{dv} b(X(s) + v(\tilde{X}(s) - X(s)), Y(s)) \right] dv ds \\ &= \int_0^t \left[ \int_0^1 b_x(X(s) + v(\tilde{X}(s) - X(s)), Y(s)) dv \right] Z^\Delta(s) ds, \end{aligned}$$

donde  $Z^\Delta(s)$  está definido en la **Definición 2.3.17** y  $b_x(\cdot)$  denota a la derivada parcial de  $b(\cdot)$  con respecto a  $x$ , (i.e.  $b_x = \frac{\partial}{\partial x} b$ ). Se sigue de la **Proposición 2.3.19** que para todo  $s \in [0, T]$ ,

$$\tilde{X}(s) - X(s) \rightarrow 0,$$

en probabilidad si  $\Delta \rightarrow 0$ . Lo cual implica que:

$$\int_0^1 b_x(X(s) + v(\tilde{X}(s) - X(s)), Y(s)) dv \rightarrow b_x(X(s), Y(s)), \quad (81)$$

en probabilidad si  $\Delta \rightarrow 0$ . De manera similar tenemos que:

$$\begin{aligned} & \frac{1}{\Delta} \int_0^t \left[ \sigma(\tilde{X}(s), \tilde{Y}(s)) - \sigma(\tilde{X}(s), Y(s)) \right] dw(s) \\ &= \int_0^t \left[ \int_0^1 \sigma_x(X(s) + v(\tilde{X}(s) - X(s)), Y(s)) dv \right] Z^\Delta(s) dw(s), \end{aligned}$$

y

$$\int_0^1 \sigma_x(X(s) + v(\tilde{X}(s) - X(s)), Y(s)) dv \rightarrow \sigma_x(X(s), Y(s)), \quad (82)$$

en probabilidad si  $\Delta \rightarrow 0$ . Sea  $\zeta(t) := \zeta^{x, \alpha}(t)$  la solución de:

$$\zeta(t) = 1 + \int_0^t b_x(X(s), Y(s))\zeta(s)ds + \int_0^t \sigma_x(X(s), Y(s))\zeta(s)dw(s), \quad (83)$$

donde  $b_x$  y  $\sigma_x$  denotan a las derivadas parciales de  $b$  y  $\sigma$  con respecto a  $x$ , respectivamente. Entonces (79), (82) y el Teorema 5.5.2 de [17] implican que:

$$\mathbb{E} \left[ \left| Z^\Delta(t) - \zeta(t) \right|^2 \right] \rightarrow 0 \text{ cuando } \Delta \rightarrow 0, \quad (84)$$

y  $\zeta(t) = \zeta^{x,\alpha}(t)$  es continua en media cuadrada con respecto a  $x$ . Por lo tanto  $\frac{\partial}{\partial x} X(t)$  existe en el sentido de media cuadrada y  $\frac{\partial}{\partial x} X(t) = \zeta(t)$ .

De manera análoga, puede demostrarse que  $\frac{\partial^2}{\partial x^2} X^{x,\alpha}(t)$  existe en el sentido de media cuadrada y es continua en media cuadrada con respecto a  $x$ . La demostración de este teorema ha sido finalizada.

Q.E.D.

**Corolario 2.3.21:** *Bajo las condiciones del Teorema 2.3.16, las derivadas en media cuadrada  $\frac{\partial}{\partial x_j} X^{x,\alpha}(t)$  y  $\frac{\partial^2}{\partial x_j \partial x_k} X^{x,\alpha}(t)$  para  $j, k = 1, \dots, n$  son continuas en media cuadrada con respecto a  $t$ .*

(La demostración de este resultado puede encontrarse en [36]).

## 2.4 Recurrencia

En esta sección se analizarán ciertas propiedades que satisface el proceso  $(X(t), Y(t))$  que satisface (28) y (29) con respecto a la recurrencia.

**Nota:** en este caso supondremos que el componente continuo del proceso, es decir  $X(t)$ , es  $r$ -dimensional, en otras palabras  $X(t) \in \mathbb{R}^r : r \geq 1$ . A diferencia de (28) y (29) donde  $X(t) \in \mathbb{S} \subset \mathbb{R}$ .

Para realizar dicho análisis se ocupará el operador  $\mathcal{L}$  definido en (36). Durante esta sección se supondrá en todo momento que los coeficientes  $\sigma(\cdot, i)$  y  $b(\cdot, i)$  satisfacen la condición local de Lipschitz y la condición de crecimiento local lineal para cada  $i \in S$  y que  $Q(\cdot)$  está acotado y es continuo. Tal y como se prueba en el **Teorema 2.1.6**, el sistema (28)-(29) posee solución única, a la cual denotaremos por  $(X^{x,\alpha}(t), Y^{x,\alpha}(t))$ .

Para estudiar la recurrencia de  $(X(t), Y(t))$ , asumimos que se satisface la siguiente condición **(A1.5.5)** por lo que resta de la sección:

**(A1.5.5)**

El operador  $\mathcal{L}$  satisface las siguientes condiciones. Para todo  $i \in S$ ,  $a(x, i) = [a_{jk}(x, i)]$  es una matriz simétrica y satisface:

$$\kappa_1 |\epsilon|^2 \leq \epsilon^* a(x, i) \epsilon \leq \kappa_1^{-1} |\epsilon|^2, \text{ para todo } \epsilon \in \mathbb{R}^r, \quad (85)$$

para alguna constante  $\kappa_1 \in (0, 1]$  para todo  $x \in \mathbb{R}^r$ .  $Q(\cdot) : \mathbb{R}^r \rightarrow \mathbb{R}^{N \times N}$  es una función matricial continua y acotada. Además  $Q(x)$  es irreducible para todo  $x \in \mathbb{R}^r$ .

Ahora, procederemos a definir los conceptos de *recurrencia*, *recurrencia positiva* y *recurrencia nula*. [**Nota:** más adelante volveremos a definir los conceptos de recurrencia, recurrencia positiva y recurrencia nula, la diferencia está en que la siguiente definición estará en términos de las *ecuaciones diferenciales asociadas a algunos operadores del proceso*].

**Definición 2.4.1** Para cualesquiera  $D \subset \mathbb{R}^r, J \subset S$  y  $U = D \times J \subset \mathbb{R}^r \times S$  denotaremos:

$$\begin{aligned}\tau_U &:= \inf\{t \geq 0 : (X(t), Y(t)) \notin U\}, \\ \sigma_U &:= \inf\{t \geq 0 : (X(t), Y(t)) \in U\}.\end{aligned}\tag{86}$$

\end{equation}

En particular, si  $U = D \times S$  es un ‘cilindro’, es decir  $D \times S = \{(v, i) : \|v\|_2 \leq \epsilon, i \in J | \epsilon > 0\}$ , tendremos que:

$$\begin{aligned}\tau_U &:= \inf\{t \geq 0 : X(t) \notin D\} \\ \sigma_U &:= \inf\{t \geq 0 : X(t) \in D\}.\end{aligned}\tag{87}$$

Definimos la recurrencia y recurrencia positiva de la siguiente manera:

- *Recurrencia y Transitividad.* Para  $U = D \times J \subset \mathbb{R}^r \times S$ , donde  $D$  es un conjunto abierto con cerradura compacta, sea:

$$\sigma_U^{x,\alpha} = \inf\{t \geq 0 : (X^{x,\alpha}(t), Y^{x,\alpha}(t)) \in U\}.$$

El proceso regular  $(X^{x,\alpha}(\cdot), Y^{x,\alpha}(\cdot))$  es *recurrente* con respecto a  $U$  si  $\mathbb{P}[\sigma_U^{x,\alpha} < \infty] = 1$  para todo  $(x, \alpha) \in D^c \times S$ . De otra forma, diremos que el proceso es *transitivo*.

- *Recurrencia Positiva y Recurrencia Nula.* Un proceso recurrente con tiempo medio de recurrencia finito para algún  $U = D \times J \subset \mathbb{R}^r \times S$ , donde  $D$  es un conjunto abierto con cerradura compacta es *recurrente positivo* con respecto a  $U$ . De otra forma, diremos que el proceso es *recurrente nulo* con respecto a  $U$ .

Mostraremos algunos resultados necesarios para proceder. El siguiente teorema nos asegura que bajo las condiciones **(A1.5.5)**, el proceso  $(X(t), Y(t))$  saldrá de cualquier ‘cilindro’ acotado con tiempo medio finito.

**Teorema 2.4.2:** Sea  $D \subseteq \mathbb{R}^r$  un conjunto abierto no vacío con cerradura compacta  $\bar{D}$ . Sea  $\tau_D := \inf\{t \geq 0 : X(t) \notin D\}$ . Entonces:

$$\mathbb{E}_{(x,i)}[\tau_D] < \infty, \text{ para todo } (x, i) \in D \times S.\tag{88}$$

**Demostración:**

Primero, nótese que por la condición (85) tendremos que:

$$\kappa_1 \leq a_{11}(x, i) \leq \kappa_1^{-1}, \text{ para todo } (x, i) \in D \times S.\tag{89}$$

Para cada  $i \in S$ , considérese:

$$W(x, i) = k - (x_1 + \beta)^c,$$

donde las constantes  $k, c$  (con  $c \geq 2$ ) y  $\beta$  serán determinadas después, además  $x_1 = e_1^*x$  es la primera componente de  $x$  donde  $e_1 = [1, 0, \dots, 0]^*$ . Calculando directamente tendremos que:

$$\mathcal{L}W(x, i) = -c(x_1 + \beta)^{c-2} \left[ b_1(x, i)(x_1 + \beta) + \frac{c-1}{2} a_{11}(x, i) \right].$$

Ajustando

$$c = \frac{2}{\kappa_1} \left( \sup_{(x,i) \in \bar{D} \times S} |b_1(x,i)(x_1 + \beta)| + 1 \right) + 1.$$

Entonces se sigue de (89) que

$$\frac{c-1}{2} a_{11}(x,i) + b_1(x,i)(x_1 + \beta) \geq \frac{c-1}{2} \kappa_1 - \sup_{(x,i) \in \bar{D} \times S} |b_1(x,i)(x_1 + \beta)| \geq 1.$$

Mientras que, dado que  $x \in D \subset \bar{D}$  donde  $\bar{D}$  es compacto, podremos escoger  $\beta$  tal que  $1 \leq x_1 + \beta \leq M$  para todo  $x \in D$ , donde  $M$  es una constante positiva. Por lo tanto  $(x_1 + \beta)^{c-2} \geq 1^{c-2} = 1$ . Finalmente seleccionamos  $k$  suficientemente grande de forma que  $W(x,i) = k - (x_1 + \beta)^c > 0$  para todo  $(x,i) \in D \times S$ . Por lo tanto  $W(x,i), i \in S$  son funciones de Liapunov que satisfacen:

$$\mathcal{L}W(x,i) \leq -c, \text{ para todo } (x,i) \in D \times S. \quad (90)$$

Sea  $\tau_D(t) = t \wedge \tau_D := \min(t, \tau_D)$ . Se sigue de la fórmula de Dynkin (véase **Teorema 1.4.16**) y de (90) que:

$$\mathbb{E}_{(x,i)} [W(X(\tau_D(t)), Y(\tau_D(t))) - W(x,i)] = \mathbb{E}_{(x,i)} \left[ \int_0^{\tau_D(t)} \mathcal{L}W(X(u), Y(u)) du \right] \leq -c \mathbb{E}_{(x,i)} [\tau_D(t)].$$

Dado que  $W(\cdot, \cdot)$  es no-negativa, se tiene que:

$$\mathbb{E}_{(x,i)} [\tau_D(t)] \leq \frac{1}{c} W(x,i). \quad (91)$$

Debido a que

$$\mathbb{E}_{(x,i)} [\tau_D(t)] = \mathbb{E}_{(x,i)} [\tau_D \mathbb{1}_{\tau_D \leq t}] + \mathbb{E}_{(x,i)} [t \mathbb{1}_{\tau_D > t}],$$

tendremos, gracias a (91), que

$$t \mathbb{P}_{(x,i)} [\tau_D > t] \leq \frac{1}{c} W(x,i).$$

Haciendo  $t \rightarrow \infty$ , tendremos que

$$\mathbb{P}_{(x,i)} [\tau_D = \infty] = \mathbb{P}_{(x,i)} [\tau_D \leq \infty] = 1.$$

Esto implica que  $\tau_D(t) \rightarrow \tau_D$  casi seguramente (c.s)  $\mathbb{P}_{(x,i)}$  cuando  $t \rightarrow \infty$ . Ahora, al aplicar el Lema de Fatou (véase **Lema 1.4.14**), si  $t \rightarrow \infty$  tendremos que:

$$\mathbb{E}_{(x,i)} [\tau_D] \leq \frac{1}{c} W(x,i) < \infty,$$

tal como se deseaba.

Q.E.D.

**Lema 2.4.3:** Para todo  $U = D \times J \subset \mathbb{R}^r \times S$ , donde  $D \subseteq \mathbb{R}^r$  un dominio no vacío, las funciones:

$$f(x, i) = \mathbb{P}_{(x, i)}[\tau_U < \infty] \text{ y } g(x, i) = \mathbb{E}_{(x, i)}[\phi(X(\tau_U), Y(\tau_U))],$$

son  $\mathcal{L}$ -armónicas en  $U$ , donde  $\phi$  es cualquier función acotada y Borel medible en  $\partial D \times S$ .

**Demostración:**

Fijando  $(x, i) \in U$ . Considérese cualquier  $V = \tilde{D} \times \tilde{J} \subset U$  tal que  $x \in \tilde{D} \subset D$  e  $i \in \tilde{J} \subset J$ . Entonces se sigue de la propiedad fuerte de Markov que

$$\begin{aligned} f(x, i) &= \mathbb{E}_{(x, i)}[\mathbb{1}_{\{\tau_U < \infty\}}] = \mathbb{E}_{(x, i)}\left[\mathbb{E}_{(x, i)}[\mathbb{1}_{\{\tau_U < \infty\}}] \Big| \mathcal{F}_{\tau_V}\right] \\ &= \mathbb{E}_{(x, i)}\left[\mathbb{E}_{(X(\tau_V), Y(\tau_V))}[\mathbb{1}_{\{\tau_U < \infty\}}]\right] = \mathbb{E}_{(x, i)}[f(X(\tau_V), Y(\tau_V))]. \end{aligned}$$

Con lo cual se prueba que  $f$  es  $\mathcal{L}$ -armónica en  $U$ . Con un argumento muy similar al anterior se prueba que  $g$  es  $\mathcal{L}$ -armónica en  $U$ .

Q.E.D.

**Lema 2.4.4:** *Supóngase (A1.5.5). Sea  $U = D \times J \subset \mathbb{R}^r \times S$  y  $f : U \rightarrow \mathbb{R}$ , donde  $D \subseteq \mathbb{R}^r$  es un dominio no vacío. Entonces:*

$$\mathcal{L}f(x, i) = 0 \text{ para todo } (x, i) \in U, \quad (92)$$

si y solo si  $f$  es  $\mathcal{L}$ -armónica en  $U$ . Además, supóngase que  $\partial D$  es suficientemente suave,  $\bar{D}$  es compacto y que  $\phi(\cdot, i)$  es una función continua arbitraria en  $\partial D$  para todo  $i \in S$ . Entonces:

$$u(x, i) := \mathbb{E}_{(x, i)}[\phi(X(\tau_U), Y(\tau_U))] \quad (93)$$

es la única solución de la ecuación diferencial (92) con condiciones de frontera:

$$\lim_{x \rightarrow x_0, x \in D} u(x, i) = \phi(x_0, i) \text{ para todo } (x_0, i) \in \partial D \times S. \quad (94)$$

(La demostración de este resultado puede encontrarse en [36]).

Ocupando un argumento similar al de la prueba del **Lema 2.4.4** es que se demuestra el siguiente resultado:

**Lema 2.4.5:** *Sea  $U = D \times S \subset \mathbb{R}^r \times S$ , donde  $D \subset \mathbb{R}^r$  es un subconjunto no vacío, abierto con cerradura compacta. Supóngase que  $g(\cdot, i) \in C_b(\bar{D})$  y  $f(\cdot, \cdot) : \bar{D} \times S \rightarrow \mathbb{R}$ . Entonces  $f$  es solución del problema de frontera:*

$$\begin{cases} \mathcal{L}f(x, i) = -g(x, i), & (x, i) \in D \times S, \\ f(x, i) = 0, & (x, i) \in \partial D \times S \end{cases}$$

si y solo si:

$$f(x, i) = \mathbb{E}_{(x, i)}\left[\int_0^{\tau_U} g(X(t), Y(t))dt\right], \text{ para todo } (x, i) \in D \times S.$$

Por medio de **Lema 2.4.3** y **Lema 2.4.4** es que procederemos a probar que si el proceso  $(X(t), Y(t))$  es recurrente (o recurrente positivo) con respecto a algún ‘cilindro’  $D \times S \subset \mathbb{R}^r \times S$ , entonces es recurrente (o recurrente positivo) con respecto a cualquier ‘cilindro’  $E \times S \subset \mathbb{R}^r \times S$ , donde  $D$  es un dominio no vacío en  $\mathbb{R}^r$  con cerradura compacta.



**Lema 2.4.6:** Sea  $D$  un dominio no vacío en  $\mathbb{R}^r$  con cerradura compacta. Supóngase que

$$\mathbb{P}_{(x,i)}[\sigma_D < \infty] = 1 \text{ para todo } (x, i) \in D^c \times S. \quad (95)$$

Entonces, para todo conjunto abierto no vacío  $E \subset \mathbb{R}^r$  tenemos que:

$$\mathbb{P}_{(x,i)}[\sigma_E < \infty] = 1 \text{ para todo } (x, i) \in E^c \times S.$$

(La demostración de este resultado puede encontrarse en [36]).

**Lema 2.4.7:** Sea  $D$  un dominio no vacío en  $\mathbb{R}^r$  con cerradura compacta. Supóngase que

$$\mathbb{E}_{(x,i)}[\sigma_D] < \infty \text{ para todo } (x, i) \in D^c \times S. \quad (96)$$

Entonces, para todo conjunto abierto no vacío  $E \subset \mathbb{R}^r$  tenemos que:

$$\mathbb{E}_{(x,i)}[\sigma_E] < \infty \text{ para todo } (x, i) \in E^c \times S.$$

(La demostración de este resultado puede encontrarse en [36]).

El siguiente lema muestra que si el proceso  $(X(t), Y(t))$  alcanza el ‘cilindro’  $D \times S$  en tiempo finito casi seguramente  $\mathbb{P}_{(x,i)}$ , entonces visitará el conjunto  $D \times \{l\}$  en tiempo finito casi seguramente  $\mathbb{P}_{(x,i)}$  para todo  $l \in S$ .

**Lema 2.4.8:** Sea  $D$  un dominio no vacío en  $\mathbb{R}^r$  con cerradura compacta. Supóngase que

$$\mathbb{P}_{(x,i)}[\sigma_D < \infty] = 1 \text{ para todo } (x, i) \in D^c \times S \quad (97)$$

Entonces tenemos que:

$$\mathbb{P}_{(x,i)}[\sigma_{D \times \{l\}} < \infty] = 1 \text{ para todo } l \in S. \quad (98)$$

(La demostración de este resultado puede encontrarse en [36]).

Ocupando el **Lema 2.4.8** es que se puede demostrar que si el proceso  $(X(t), Y(t))$  es recurrente positivo con respecto a un ‘cilindro’  $D \times S$ , entonces es recurrente positivo con respecto al conjunto  $D \times \{l\} \subset \mathbb{R}^r \times S$ .

**Lema 2.4.9:** Sea  $D$  un dominio no vacío en  $\mathbb{R}^r$  con cerradura compacta. Supóngase que

$$\mathbb{E}_{(x,i)}[\sigma_D] < \infty \text{ para todo } (x, i) \in D^c \times S. \quad (99)$$

Entonces, para todo  $(x, i) \in \mathbb{R}^r \times S$  tenemos que:

$$\mathbb{E}_{(x,i)}[\sigma_{D \times \{l\}}] < \infty \text{ para todo } l \in S. \quad (100)$$

(La demostración de este resultado puede encontrarse en [36]).

Resumiremos los anteriores lemas en el siguiente teorema.

**Teorema 2.4.10:** Supóngase que se satisface **(A1.5.5)**. Entonces se satisfacen los siguientes resultados:

- El proceso  $(X(t), Y(t))$  es recurrente (o recurrente positivo) con respecto a  $D \times S$  si y solo si es recurrente (o recurrente positivo) con respecto a  $D \times \{l\}$ , donde  $D \in \mathbb{R}^r$  es un conjunto abierto no vacío con cerradura compacta y  $l \in S$ .

- Si el proceso  $(X(t), Y(t))$  es recurrente (o recurrente positivo) con respecto a  $U = D \times S$ , donde  $D \in \mathbb{R}^r$  es un conjunto abierto no vacío con cerradura compacta, entonces es recurrente (o recurrente positivo) con respecto a cualquier  $\hat{U} = \hat{D} \times S$ , donde  $\hat{D} \subset \mathbb{R}^r$  es un conjunto abierto no vacío.

Nótese que:

- La **recurrencia** es una propiedad independiente de la región escogida; por lo tanto, el proceso  $(X(t), Y(t))$  con generador  $\mathcal{L}$  asociado que satisface **(A1.5.5)** se le llama **recurrente** si es recurrente con respecto a algún  $U = D \times \{l\}$  donde  $D \in \mathbb{R}^r$  es un conjunto abierto no vacío y  $l \in S$ ; de otra forma le llamaremos **transitivo**.
- Por ende, llamamos al proceso  $(X(t), Y(t))$  **positivo recurrente** si es positivo recurrente con respecto a algún dominio acotado  $U = D \times \{l\} \subset \mathbb{R}^r \times S$ ; de otra forma le llamaremos **nulo recurrente**.

Para poder estudiar la recurrencia del proceso  $(X(t), Y(t))$ , primero analizaremos el siguiente criterio basado en la existencia de ciertas **funciones de Liapunov**.

**Teorema 2.4.11:** *Suponga la existencia de un conjunto abierto no vacío  $D \subset \mathbb{R}^r$  bajo el cual existe  $V(\cdot, \cdot) : D^c \times S \rightarrow \mathbb{R}^+$  que satisface:*

$$\begin{aligned} V_n &:= \inf_{|x| \geq n, i \in S} V(x, i), \quad \lim_{n \rightarrow \infty} V_n = \infty, \\ \mathcal{L}V(x, i) &\leq 0, \quad \text{para todo } (x, i) \in D^c \times S. \end{aligned} \tag{101}$$

Entonces el proceso  $(X(t), Y(t))$  es recurrente.

**Demostración:**

Sea  $(x, \alpha) \in D^c \times S$  fijo. Definimos:

$$\sigma_D = \sigma_D^{x, \alpha} := \inf\{t \geq 0 : X^{x, \alpha}(t) \in D\},$$

y

$$\sigma_D^{(n)}(t) := \sigma_D \wedge t \wedge \beta_n,$$

donde  $\beta_n = \inf\{t > 0 : |X^{(n)}(t)| \geq n\}$  como se definió en (44). Entonces se sigue de la fórmula de Dynkin (véase **Teorema 1.4.16**) que:

$$\mathbb{E} \left[ V \left( X \left( \sigma_D^{(n)}(t) \right), Y \left( \sigma_D^{(n)}(t) \right) \right) - V(x, \alpha) \right] = \mathbb{E} \left[ \int_0^{\sigma_D^{(n)}(t)} \mathcal{L}V(X(u), Y(u)) du \right] \leq 0.$$

Por ende,

$$\mathbb{E} \left[ V \left( X \left( \sigma_D^{(n)}(t) \right), Y \left( \sigma_D^{(n)}(t) \right) \right) \right] \leq V(x, \alpha).$$

Nótese que si  $t \rightarrow \infty$ ,  $\sigma_D \wedge t \wedge \beta_n \rightarrow \sigma_D \wedge \beta_n$  c.s. Por lo que el lema de Fatou (véase **Lema 1.4.14**) implica que:

$$\mathbb{E} \left[ V \left( X \left( \sigma_D \wedge \beta_n \right), Y \left( \sigma_D \wedge \beta_n \right) \right) \right] \leq V(x, \alpha).$$

Por lo cual se tiene que

$$V(x, \alpha) \geq \mathbb{E}[V(X(\sigma_D \wedge \beta_n), Y(\sigma_D \wedge \beta_n))] \geq \mathbb{E}[V(X(\beta_n), Y(\beta_n)) \mathbb{1}_{\{\beta_n < \sigma_D\}}] \geq V_n \mathbb{P}[\beta_n < \sigma_D].$$

Se sigue de (101) que si  $n \rightarrow \infty$ ,

$$\mathbb{P}[\beta_n < \sigma_D] \leq \frac{V(x, \alpha)}{V_n} \rightarrow 0.$$

Nótese que:

$$\mathbb{P}[\sigma_D = \infty] \leq \mathbb{P}[\beta_n < \sigma_D].$$

Por lo tanto  $\mathbb{P}[\sigma_D < \infty] = 1$ , tal y como se deseaba.

Q.E.D.

Construir una función de Liapunov por lo general es muy difícil. El siguiente teorema muestra una forma menos complicada de demostrar la recurrencia del proceso  $(X(t), Y(t))$ .

**Teorema 2.4.12:** *Al menos una de las siguientes condiciones satisfechas implica que el proceso  $(X(t), Y(t))$  es recurrente.*

(i) *Existen constantes  $\gamma > 0$  y  $c_i \in \mathbb{R}$  donde  $i \in S$  tal que para todo  $(x, i) \in \{c \in \mathbb{R}^r : |c| \geq 1\} \times S$ ,*

$$\frac{x^*b(x, i)}{|x|^2} + \frac{\text{tr}(a(x, i))}{|x|^2} + (\gamma - 2) \frac{x^*a(x, i)x}{2|x|^4} - \frac{1}{k - \gamma c_i} \sum_{j=1}^N q_{ij}(x)c_j \leq 0, \quad (102)$$

donde  $k$  es una constante suficientemente grande de forma que  $k - \gamma c_i > 0$  para todo  $i \in S$ .

(ii) *Existe una constante positiva  $\gamma$  y matrices definidas positivas  $P_i$  para  $i \in S$  tales que para  $(x, i) \in \{c \in \mathbb{R}^r : |c| \geq 1\} \times S$ ,*

$$\frac{x^*P_i b(x, i)}{x^*P_i x} + \frac{\text{tr}(\sigma^*(x, i)P_i \sigma(x, i))}{2x^*P_i x} + (\gamma - 2) \frac{|\sigma^*(x, i)P_i x|^2}{|x^*P_i x|^2} + \sum_{j=1}^N q_{ij}(x) \frac{|x^*P_j x|^{\frac{\gamma}{2}}}{|x^*P_i x|^{\frac{\gamma}{2}}} \leq 0. \quad (103)$$

**Demostración:**

Para todo  $i \in S$ , defínase la función de Liapunov:

$$V(x, i) = (k - \gamma c_i)|x|^\gamma.$$

Entonces un cálculo directo nos muestra que para todo  $x \neq 0$ , tendremos que

$$\begin{aligned} \nabla V(x, i) &= (k - \gamma c_i)\gamma|x|^{\gamma-2}x, \\ \nabla^2 V(x, i) &= (k - \gamma c_i)\gamma [|x|^{\gamma-2}I + (\gamma - 2)|x|^{\gamma-4}xx^*]. \end{aligned}$$

Por lo tanto se sigue que:

$$\mathcal{L}V(x, i) = (k - \gamma c_i) \gamma |x|^\gamma \left[ \frac{x^* b(x, i)}{|x|^2} + \frac{\text{tr}(a(x, i))}{|x|^2} - (\gamma - 2) \frac{x^* a(x, i) x}{|x|^4} - \frac{1}{k - \gamma c_i} \sum_{j=1}^N q_{ij}(x) c_j \right].$$

Por lo tanto el **Teorema 2.4.11** implica la conclusión deseada de la condición (i).

Para probar el caso de la condición (ii) se ocupa un argumento similar al del caso de la condición (i) al considerar:

$$W(x, i) = (x^* P_i x)^{\frac{\gamma}{2}} \text{ para todo } (x, i) \in \{x \in \mathbb{R}^r : |x| \geq 1\} \times S,$$

y verificando que

$$\begin{aligned} \nabla W(x, i) &= \gamma (x^* P_i x)^{\frac{\gamma-2}{2}} P_i x, \\ \nabla^2 W(x, i) &= \gamma (x^* P_i x)^{\frac{\gamma-2}{2}} P_i + \gamma(\gamma-2) (x^* P_i x)^{\frac{\gamma-4}{2}} P_i x x^* P_i. \end{aligned}$$

Q.E.D.

**Lema 2.4.13:** *Si existe algún  $(x_0, l) \in \mathbb{R}^r \times S$  tal que para cualquier  $\epsilon > 0$ ,*

$$\mathbb{P}_{(x_0, l)} [(X(t_n), Y(t_n)) \in B(x_0, \epsilon) \times \{l\}, \text{ para alguna sucesión } t_n \uparrow \infty] = 1, \quad (104)$$

*entonces para cualquier  $U := D \times \{j\} \subset \mathbb{R}^r \times S$ , donde  $D$  es un dominio acotado y no vacío,  $j \in S$ , tenemos que:*

$$\mathbb{P}_{(x_0, l)} [\sigma_U < \infty] = 1.$$

*En particular, si (104) es cierto para cualquier  $(x, i) \in \mathbb{R}^r \times S$ , entonces el proceso  $(X(t), Y(t))$  es recurrente.*

**Demostración:**

Basta con considerar dos casos: cuando  $x_0 \notin D$  y cuando  $x_0 \in D$  con  $j \neq l$ . En el primer caso puede adaptarse la prueba del **Lema 2.4.6** de forma que se demuestre que  $\mathbb{P}_{(x_0, l)} [\sigma_D < \infty = 1]$ . Entonces, en virtud del **Lema 2.4.8**, se sigue que  $\mathbb{P}_{(x_0, l)} [\sigma_{D \times \{j\}} < \infty = 1]$ . El segundo caso se sigue de una ligera modificación del argumento ocupado en la prueba del **Lema 2.4.8**.

Q.E.D.

**Lema 2.4.14:** *Si el proceso  $(X(t), Y(t))$  es recurrente, entonces para cualquier  $(x, i) \in \mathbb{R}^r \times S$  y  $\epsilon > 0$  tenemos que:*

$$\mathbb{P}_{(x, i)} [(X(t_n), Y(t_n)) \in B(x, \epsilon) \times \{i\}, \text{ para alguna sucesión } t_n \uparrow \infty] = 1. \quad (105)$$

**Demostración:**

Denotemos  $B = B(x, \epsilon)$ ,  $B_1 = B(x, \frac{\epsilon}{2})$ , y  $B_2 = B(x, 2\epsilon)$ . Defínase una sucesión de tiempos de paro por:

$$\eta_1 := \inf\{t \geq 0 : X(t) \notin B_2\};$$

y para todo  $n = 1, 2, \dots$ ,

$$\begin{aligned}\eta_{2n} &:= \inf\{t \geq \eta_{2n-1} : (X(t), Y(t)) \in B_1 \times \{i\}\}; \\ \eta_{2n+1} &:= \inf\{t \geq \eta_{2n} : X(t) \notin B_2\}.\end{aligned}$$

Nótese que el proceso  $(X(t), Y(t))$  es recurrente, en particular, es recurrente con respecto a  $B_1 \times \{i\}$ . Esto, junto con el **Teorema 2.4.2**, implica que  $\eta_n < \infty$  c.s. en  $\mathbb{P}_{(x,i)}$ . Por lo tanto se sigue (105).

Q.E.D.

Al combinar los **Lema 2.4.13** y **Lema 2.4.14** se obtiene el siguiente resultado:

**Teorema 2.4.15:** *El proceso  $(X(t), Y(t))$  es recurrente si y solo si todo punto  $(x, i) \in \mathbb{R}^r \times S$  es recurrente en el sentido de que para todo  $\epsilon > 0$ ,*

$$\mathbb{P}_{(x,i)} [(X(t_n), Y(t_n)) \in B(x, \epsilon) \times \{i\}, \text{ para alguna sucesión } t_n \uparrow \infty] = 1.$$

El **Teorema 2.4.15** nos permite ajustar otro criterio para la recurrencia en términos del tiempo de estadía. Para lograrlo, para todo  $U := D \times J \subset \mathbb{R}^r \times S$  y cualquier  $\lambda \geq 0$ , definimos:

$$R_\lambda(x, i, U) := \mathbb{E}_{(x,i)} \left[ \int_0^\infty e^{-\lambda t} \mathbb{1}_U(X(t), Y(t)) dt \right]. \quad (106)$$

En particular,  $R_0(x, i, U)$  denota al tiempo medio de estadía del proceso  $(X^{x,i}(t), Y^{x,i}(t))$  en el dominio  $U$ .

**Proposición 2.4.16:** *Supóngase (A1.5.5). Si para algún punto  $(x_0, l) \in \mathbb{R}^r \times S$  y cualquier  $\rho > 0$ ,*

$$R_0(x_0, l, B(x_0, \rho) \times \{l\}) = \infty, \quad (107)$$

*entonces  $(x_0, l)$  es un punto recurrente, es decir, para todo  $\rho > 0$ :*

$$\mathbb{P}_{(x_0,l)} [(X(t_n), Y(t_n)) \in B(x_0, \epsilon) \times \{l\}, \text{ para alguna sucesión } t_n \uparrow \infty] = 1. \quad (108)$$

*En particular, si (107) es válido para cualquier  $(x, i) \in \mathbb{R}^r \times S$ , entonces el proceso  $(X(t), Y(t))$  es recurrente.*

**Demostración:**

En vista del **Teorema 2.4.15**, basta con demostrar la primera afirmación. Denótese por  $B = B(x_0, \rho)$  y a  $U = B \times \{l\}$ . Defínase para toda  $T > 0$  y cualquier  $0 < \epsilon < \rho$  la sucesión de tiempos de paro  $\zeta_0 := 0$ , para todo  $n \geq 1$ ,

$$\zeta_n := \inf\{t > \zeta_{n-1} + T : (X(t); Y(t)) \in B(x_0, \epsilon) \times \{l\}\}.$$

(Se ocupa la convención  $\inf\{\emptyset\} = \infty$ ). En virtud de la propiedad fuerte de Markov, tenemos que

$$\sup_{x \in B(x_0, \epsilon)} \mathbb{P}_{(x,l)} [\zeta_n < \infty] \leq \left( \sup_{x \in B(x_0, \epsilon)} \mathbb{P}_{(x,l)} [\zeta_1 < \infty] \right)^n. \quad (109)$$

Por lo que

$$\begin{aligned}R_0(x_0, l, U) &= \sum_{n=0}^{\infty} \mathbb{E}_{(x_0,l)} \left[ \mathbb{1}_{\{\zeta_n < \infty\}} \int_{\zeta_n}^{\zeta_{n+1}} \mathbb{1}_U(X(t), Y(t)) dt \right] \\ &= \sum_{n=0}^{\infty} \mathbb{E}_{(x_0,l)} \left[ \mathbb{1}_{\{\zeta_n < \infty\}} \mathbb{E}_{(X(\zeta_n), Y(\zeta_n))} \left[ \int_0^{\zeta_1} \mathbb{1}_U(X(t), Y(t)) dt \right] \right].\end{aligned}$$

Nótese que

$$\int_0^{\zeta_1} \mathbb{1}_U(X(t), Y(t)) dt = \int_0^T \mathbb{1}_U(X(t), Y(t)) dt + \int_T^{\zeta_1} \mathbb{1}_U(X(t), Y(t)) dt \leq T.$$

Se sigue de (109) que

$$R_0(x_0, l, U) \leq T \sum_{n=0}^{\infty} \mathbb{P}_{(x_0, l)} [\zeta_n < \infty] \leq T \sum_{n=0}^{\infty} \left( \sup_{x \in B(x_0, \epsilon)} \mathbb{P}_{(x, l)} [\zeta_1 < \infty] \right)^n.$$

Dado que  $R_0(x_0, l, U) = \infty$ , se tiene que

$$\sup_{x \in B(x_0, \epsilon)} \mathbb{P}_{(x, l)} [\sigma_{B(x_0, \epsilon) \times \{l\}}^T < \infty] = \sup_{x \in B(x_0, \epsilon)} \mathbb{P}_{(x, l)} [\zeta_1 < \infty] = 1,$$

donde  $\sigma_{B(x_0, \epsilon) \times \{l\}}^T := \inf\{t > T : (X(t), Y(t)) \in B(x_0, \epsilon) \times \{l\}\}$ . Dado que  $B(x_0, \epsilon) \subset B$ ,

$$\mathbb{P}_{(x, l)} [\sigma_{B(x_0, \epsilon) \times \{l\}}^T < \infty] \leq \mathbb{P}_{(x, l)} [\sigma_{B \times \{l\}}^T < \infty],$$

y  $\sup_{x \in B(x_0, \epsilon)} \mathbb{P}_{(x, l)} [\sigma_{B \times \{l\}}^T < \infty] = 1$ . Finalmente, dado que  $(X(t), Y(t))$  satisface la propiedad fuerte de Feller, se obtiene que si hacemos  $\epsilon \rightarrow 0$ , entonces:

$$\mathbb{P}_{(x, l)} [\sigma_{B \times \{l\}}^T < \infty] = 1,$$

para todo  $T > 0$ . Por lo que se sigue (108). Con lo cual se concluye la prueba de la proposición.

Q.E.D.

Primero aclararemos que si el proceso  $(X(t), Y(t))$  es transitivo, entonces la norma del componente continuo  $|X(t)| \rightarrow \infty$  casi seguramente si  $t \rightarrow \infty$  y viceversa. Luego brindaremos un par de criterios de transitividad en términos del tiempo medio de estadía y de las funciones de Liapunov, respectivamente. Si los coeficientes  $\sigma(x, i)$  y  $b(x, i)$  para todo  $i \in S$  son linealizables en  $x$ , entonces obtenemos condiciones para la transitividad fáciles de verificar.

**Teorema 2.4.17:** *El proceso  $(X(t), Y(t))$  es transitivo si y solo si:*

$$\lim_{t \rightarrow \infty} |X(t)| = \infty \text{ c.s. en } \mathbb{P}_{(x, i)} \text{ para todo } (x, i) \in \mathbb{R}^r \times S.$$

**Demostración:**

La prueba de este resultado se sigue del argumento ocupado por Bhattacharya en [5]:

*Suficiencia*

Basta con ocupar un argumento por contradicción.

*Necesidad*

Supóngase que el proceso  $(X(t), Y(t))$  es transitivo. Fijando  $(x, \alpha) \in \mathbb{R}^r \times S$ , sea  $\rho \in \mathbb{R}$  lo suficientemente grande de forma que  $\rho > 1 \vee |x|$ . El proceso  $(X(t), Y(t))$  es transitivo, por lo que es transitivo con respecto al ‘cilindro’  $B(0, \rho) \times S$ . Por lo que existe algún  $(y_0, j_0) \in (\mathbb{R}^r - B(0, \rho)) \times S$  tal que

$$\mathbb{P}_{(y_0, j_0)} [\sigma_{B(0, \rho)} < \infty] < 1. \tag{110}$$

Supóngase que  $|y_0| = r_0 > \rho$ . Por lo que, en virtud de **Lema 2.4.3**, la función  $(y, j) \mapsto \mathbb{P}_{(y, j)} [\sigma_{B(0, \rho)} < \infty]$  es  $\mathcal{L}$ -armónica. Por lo que se sigue del principio del máximo (véase **Lema 1.4.13** y [18]) y de (110) que

$$\sup_{|y|=r_0, j \in S} \mathbb{P}_{(y, j)} [\sigma_{B(0, \rho)} < \infty] = \delta < 1. \quad (111)$$

Entonces ocupando un argumento estándar (véase Teorema 3.2 de [5]), puede probarse que

$$\mathbb{P}_{(x, \alpha)} \left[ \liminf_{t \rightarrow \infty} |X(t)| > \rho - 1 \right] = 1.$$

Dado que lo anterior se satisface para todo  $\rho > 0$ ,

$$\mathbb{P}_{(x, \alpha)} [|X(t)| \rightarrow \infty \text{ cuando } t \rightarrow \infty] = 1,$$

tal como se tenía deseado.

Q.E.D.

Enunciamos un criterio de transitividad bajo ciertas condiciones.

**Proposición 2.4.18:** *Supóngase que se satisfacen las siguientes condiciones:*

- Para  $i = 1, 2, \dots, N$ , los coeficientes  $b(\cdot, i)$ ,  $\sigma(\cdot, i)$  y  $Q(\cdot)$  son Hölder continuos con exponente  $0 < \gamma \leq 1$  (véase **Definición 1.4.3**).
- $Q(x)$  es irreducible para todo  $x \in \mathbb{R}^r$  (es decir, que el proceso  $Y(\cdot)$  cuya matriz de intensidad de transición es  $Q(x)$  satisface que todos los estados  $S = \{1, \dots, N\}$  son alcanzables desde cualquier estado  $i \in S$ ).
- Para todo  $i \in S$ ,  $a(x, i) = \sigma(x, i)\sigma^t(x, i)$  es una matriz simétrica y para alguna constante  $\kappa \in \mathbb{R}_+$  y para todo  $x \in \mathbb{R}^r$  satisface:

$$\langle a(x, i)\epsilon, \epsilon \rangle \geq \kappa|\epsilon|^2, \text{ para todo } \epsilon \in \mathbb{R}^r, \quad (112)$$

Si existe un conjunto  $U := D \times J \subset \mathbb{R}^r \times S$  que contiene el punto  $(x_0, l)$ , donde  $D$  es un conjunto abierto no vacío y acotado tal que  $R_0(x_0, l, U) < \infty$  (106), entonces el proceso  $(X(t), Y(t))$  es transitivo.

**Demostración:**

En virtud del **Teorema 2.4.15**, basta con demostrar que el punto  $(x_0, l)$  es transitivo en el sentido de que existe  $\epsilon_0 > 0$  y un tiempo finito  $T_0 > 0$  tal que

$$\mathbb{P}_{(x_0, l)} [(X(t), Y(t)) \notin B(x_0, \epsilon_0) \times \{l\}, \text{ para todo } t \geq T_0] = 1. \quad (113)$$

Claramente  $R_0(x_0, l, U) = \mathbb{E}_{(x_0, l)} \left[ \int_0^\infty \mathbb{1}_U(X(t), Y(t)) dt \right] > 0$ . Por lo que existe un  $t_1 > 0$  tal que

$$\mathbb{E}_{(x_0, l)} \left[ \int_0^{t_1} \mathbb{1}_U(X(t), Y(t)) dt \right] > 0.$$

Debido a que el proceso  $(X(t), Y(t))$  satisface la propiedad fuerte de Feller en virtud del **Teorema 2.3.12**, se concluye que existe una vecindad  $E \subset D$  de  $x_0$  tal que

$$\inf_{y \in E} \mathbb{E}_{(y,l)} \left[ \int_0^{t_1} \mathbb{1}_U(X(t), Y(t)) dt \right] > 0.$$

Sea

$$\delta := \inf_{x \in E} R_0(x, l, U) > 0. \quad (114)$$

Para cada  $T > 0$ , defínase

$$\delta_{E \times \{l\}} = \inf\{t > T : (X(t), Y(t)) \in E \times \{l\}\}. \quad (115)$$

Se sigue de la propiedad fuerte de Markov que

$$\begin{aligned} R_0(x_0, l, U) &\geq \mathbb{E}_{(x_0,l)} \left[ \int_T^\infty \mathbb{1}_U(X(t), Y(t)) dt \right] \\ &\geq \mathbb{E}_{(x_0,l)} \left[ \mathbb{1}_{\{\sigma_{E \times \{l\}}^T < \infty\}} R_0(X(\sigma_{E \times \{l\}}^T), Y(\sigma_{E \times \{l\}}^T), U) \right] \\ &\geq \inf_{x \in E} R_0(x, l, U) \mathbb{P}_{(x_0,l)} [\sigma_{E \times \{l\}}^T < \infty]. \end{aligned}$$

Por lo que se sigue de (114) y de (115) que

$$\mathbb{P}_{(x_0,l)} [\sigma_{E \times \{l\}}^T < \infty] \leq \frac{1}{\delta} \mathbb{E}_{(x_0,l)} \left[ \int_T^\infty \mathbb{1}_U(X(t), Y(t)) dt \right].$$

Para todo  $\epsilon > 0$ , en vista de la suposición  $R_0(x_0, l, U) < \infty$ , podemos escoger algún  $\tilde{T}_0 > 0$  tal que

$$\mathbb{E}_{(x_0,l)} \left[ \int_{\tilde{T}_0}^\infty \mathbb{1}_U(X(t), Y(t)) dt \right] < \delta \epsilon,$$

y así

$$\mathbb{P}_{(x_0,l)} [\sigma_{E \times \{l\}}^{\tilde{T}_0} < \infty] < \epsilon.$$

Es decir

$$\lim_{T \rightarrow \infty} \mathbb{P}_{(x_0,l)} [\sigma_{E \times \{l\}}^T < \infty] = \mathbb{P}_{(x_0,l)} [\sigma_{E \times \{l\}}^T < \infty, \text{ para todo } T > 0] = 0.$$

Por lo que existe un  $T_0 > 0$  tal que

$$\mathbb{P}_{(x_0,l)} [\sigma_{E \times \{l\}}^{T_0} = \infty] > 0. \quad (116)$$

Por lo que se sigue (113).

Q.E.D.

**Nótese** que se sigue de la **Proposición 2.4.13** y de la **Proposición 2.4.15** que el proceso  $(X(t), Y(t))$  es recurrente si y solo si para todo  $(x, i) \in \mathbb{R}^r \times S$  y  $\rho > 0$  tenemos que:



$$R_0(x, i, B(x, \rho) \times \{i\}) = \infty.$$

Ahora, obtendremos una condición suficiente para transitividad en términos de la existencia de una función de Liapunov.

**Teorema 2.4.19:** *Supóngase que existe un dominio no vacío y acotado  $D \subset \mathbb{R}^r$  y una función  $V(\cdot, \cdot) : D^c \times S \rightarrow \mathbb{R}$  que satisfice:*

$$\begin{aligned} \sup_{(x,i) \in \partial D \times S} V(x, i) &\leq 0, \\ \mathcal{L}V(x, i) &\geq 0 \text{ para todo } (x, i) \in D^c \times S, \\ \sup_{(x,i) \in D^c \times S} V(x, i) &\leq M < \infty, \\ V(y, l) &> 0 \text{ para algún } (y, l) \in D^c \times S. \end{aligned} \tag{117}$$

Entonces el proceso  $(X(t), Y(t))$  es transitivo o es no regular.

**Demostración:**

Asumiendo que el proceso  $(X(t), Y(t))$  es regular, se necesita probar que este es transitivo. Fijando  $(y, l) \in D^c \times S$  tal que  $V(y, l) > 0$ , definimos los tiempos de paro  $\sigma_D = \sigma_D^{y,l}$ ,  $\beta_n = \beta_n^{y,l}$  y  $\beta_n \wedge \sigma_D \wedge t$  como en la prueba del **Teorema 2.4.11**, con  $n > (n_0 \vee |y|)$ , donde  $n_0$  es un entero tal que  $D \subset \{x : |x| < n_0\}$  y  $\beta_n^{y,l} = \inf\{t : |X^{y,l}(t)| = n\}$ . En virtud de la fórmula de Dynkin y de (117) se tiene que:

$$\begin{aligned} &\mathbb{E} [V(X(\beta_n \wedge \sigma_D \wedge t), Y(\beta_n \wedge \sigma_D \wedge t)) - V(y, l)] \\ &= \mathbb{E} \left[ \int_0^{\beta_n \wedge \sigma_D \wedge t} \mathcal{L}V(X(u), Y(u)) du \right] \geq 0. \end{aligned}$$

Por lo tanto se sigue de (117) que:

$$\begin{aligned} V(y, l) &\leq \mathbb{E} [V(X(\sigma_D), Y(\sigma_D)) \mathbb{1}_{\{\sigma_D \leq \beta_n \wedge t\}}] + \mathbb{E} [V(X(\beta_n \wedge t), Y(\beta_n \wedge t)) \mathbb{1}_{\{\sigma_D > \beta_n \wedge t\}}] \\ &\leq M \mathbb{P} [\sigma_D > \beta_n \wedge t]. \end{aligned}$$

Sea  $A_n := \{\omega \in \Omega : \sigma_D(\omega) > \beta_n \wedge t\}$ . Entonces:

$$\mathbb{P}[A_n] \geq \frac{V(y, l)}{M}, \text{ para todo } n \geq n_0.$$

Nótese que  $\beta_n \leq \beta_{n+1}$  implica que  $A_{n+1} \subset A_n$ . Esto, combinado con la regularidad, implica que:

$$\bigcap_{n=n_0}^{\infty} A_n = \lim_{n \rightarrow \infty} A_n = \{\sigma_D > t\}.$$

Por lo tanto:

$$\mathbb{P}[\sigma_D > t] \geq \frac{V(y, l)}{M}.$$

Finalmente, haciendo  $t \rightarrow \infty$  tendremos que:

$$\mathbb{P}[\sigma_D = \infty] > 0.$$

Por lo tanto el proceso  $(X(t), Y(t))$  es transitivo.

Q.E.D.

Para continuar, nos enfocaremos en sistemas linealizables (en el componente continuo). Además de la condición **(A1.5.5)**, asumimos que la siguiente condición se satisface:

**(A1.5.6)**

Para todo  $i \in S$ , existen  $b(i), \sigma_j(i) \in \mathbb{R}^{r \times r}, j = 1, 2, \dots, d$  y  $\hat{Q} = [\hat{q}_{ij}]$  generador de una cadena de Markov a tiempo continuo  $\hat{Y}(t)$  tal que si  $|x| \rightarrow \infty$ ,

$$\begin{aligned} \frac{b(x, i)}{|x|} &= b(i) \frac{x}{|x|} + o(1), \\ \frac{\sigma(x, i)}{|x|} &= (\sigma_1(i)x, \sigma_2(i)x, \dots, \sigma_d(i)x) \frac{1}{|x|} + o(1), \\ Q(x) &= \hat{Q} + o(1), \end{aligned} \tag{118}$$

donde  $o(1) \rightarrow 0$  si  $|x| \rightarrow \infty$ . Además,  $\hat{Y}(t)$  es irreducible con distribución estacionaria denotada por  $\pi = (\pi_1, \pi_2, \dots, \pi_N) \in \mathbb{R}^{1 \times N}$ .

**Teorema 2.4.20:** *Supóngase que se satisfacen **(A1.5.5)** y **(A1.5.6)**. Si para todo  $i \in S$  se cumple que:*

$$\lambda_{\min} \left( b(i) + b^*(i) + \sum_{j=1}^d \sigma_j(i) \sigma_j^*(i) \right) - \frac{1}{2} \sum_{j=1}^d [\lambda_{\max}(\sigma_j(i) + \sigma_j^*(i))]^2 > 0, \tag{119}$$

donde  $\lambda_{\min}(\cdot)$  y  $\lambda_{\max}(\cdot)$  son las funciones de valor propio mínimo y de valor propio máximo de una matriz respectivamente, entonces el proceso  $(X(t), Y(t))$  es transitivo.

**Demostración:**

Sea  $D = \{x \in \mathbb{R}^r : |x| < k\}$ , donde  $k$  es un número positivo lo suficientemente grande. Considérese  $W(x, i) = k^\beta - |x|^\beta, (x, i) \in D^c \times S$ , donde  $\beta > 0$  es una constante lo suficientemente pequeña. Entonces tenemos que  $W(x, i) = 0$  para todo  $(x, i) \in \partial D \times S$  y  $k^\beta \geq W(x, i) > 0$  para todo  $(x, i) \in D \times S$ . Por lo que se satisfacen las condiciones (117). De igual forma tenemos que:

$$\begin{aligned} \nabla W(x, i) &= -\beta |x|^{\beta-2} x, \\ \nabla^2 W(x, i) &= -\beta [|x|^{\beta-2} I + (\beta - 2) |x|^{\beta-4} x x^*]. \end{aligned}$$

Se sigue de (118) que para todo  $(x, i) \in D^c \times S$  se tiene que:

$$\mathcal{L}W(x, i) = -\beta |x|^\beta \left\{ \frac{x^* b(i) x}{|x|^2} + \frac{1}{2} \sum_{j=1}^d \left( \frac{x^* \sigma_j^*(i) \sigma_j(i) x}{|x|^2} + (\beta - 2) \frac{(x^* \sigma_j^*(i) x)^2}{|x|^4} \right) + o(1) \right\}. \tag{120}$$

Nótese que:

$$\begin{aligned} \frac{x^* b(i) x}{|x|^2} + \frac{1}{2} \sum_{j=1}^d \frac{x^* \sigma_j^*(i) \sigma_j(i) x}{|x|^2} &= \frac{x^* \left( b^*(i) + b(i) + \sum_{j=1}^d \sigma_j^*(i) \sigma_j(i) \right) x}{2|x|^2} \\ &\geq \frac{1}{2} \lambda_{\min} \left( b^*(i) + b(i) + \sum_{j=1}^d \sigma_j^*(i) \sigma_j(i) \right). \end{aligned} \tag{121}$$

Mientras que, para cualquier matriz simétrica  $A$  con valores propios reales  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ , ocupando la transformación  $x = Uy$  donde  $U$  es una matriz real ortogonal que satisface  $U^*AU = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$  (véase Teorema 8.11 de [20]) tendremos que:

$$|x^*Ax| = |\lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2| \leq \lambda_{\max}(A)|y|^2 = \rho_A|x|^2. \quad (122)$$

Por lo tanto se sigue de (120)-(122) que para todo  $(x, i) \in D^c \times S$ ,

$$\mathcal{L}W(x, i) \geq -\frac{1}{2}\beta|x|^\beta \left\{ \lambda_{\min} \left( b(i) + b^*(i) + \sum_{j=1}^d \sigma_j^*(i)\sigma_j(i) \right) - \frac{1}{2} \sum_{j=1}^d [\lambda_{\max}(\sigma_j^*(i) + \sigma_j(i))]^2 + o(1) + O(\beta) \right\} \geq 0.$$

En el último paso se ocupó el hecho de que  $\beta > 0$  y la condición (119).

Por ende la segunda ecuación (117) del **Teorema 2.4.19** es válida. Por lo tanto el **Teorema 2.4.19** implica que el proceso  $(X(t), Y(t))$  no es recurrente.

Q.E.D.

**Teorema 2.4.21:** *Una condición necesaria y suficiente para la recurrencia positiva con respecto al dominio  $U = D \times \{l\} \subset \mathbb{R}^r \times S$  es:*

*Para todo  $i \in S$ , existe una función no negativa  $V(\cdot, i) : D^c \rightarrow \mathbb{R}$  tal que  $V(\cdot, i)$  es continua y doblemente diferenciable y que:*

$$\mathcal{L}V(x, i) = -1, \quad (x, i) \in D^c \times S. \quad (123)$$

*Sea  $u(x, i) = \mathbb{E}_{(x, i)}[\sigma_D]$ . Entonces  $u(x, i)$  es la menor solución positiva de:*

$$\begin{cases} \mathcal{L}u(x, i) = -1, & (x, i) \in D^c \times S, \\ u(x, i) = 0, & (x, i) \in \partial D \times S. \end{cases} \quad (124)$$

**Demostración:**

*Paso 1*

Probar que el proceso  $Z(t) = (X(t), Y(t))$  es positivo recurrente si existe una función no negativa  $V(\cdot, \cdot)$  que satisface las condiciones del teorema. Seleccionamos  $n_0$  como un entero positivo lo suficientemente grande de forma que  $D \subset \{|x| < n_0\}$ . Fijando  $(x, i) \in D^c \times S$ , para cualesquiera  $t > 0$  y  $n \in \mathbb{N}$  donde  $n > n_0$ , definimos:

$$\sigma_D^{(n)}(t) = \sigma_D \wedge t \wedge \beta_n,$$

donde  $\beta_n = \inf\{t : |X^{(n)}(t)| \geq n\}$  es definido como en (44) y  $\sigma_D$  es el primer tiempo de entrada de  $X(t)$  a  $D$ , es decir,  $\sigma_D = \inf\{t : X(t) \in D\}$ . Por lo que la fórmula de Dynkin y la ecuación (123) implican que:

$$\begin{aligned} & \mathbb{E}_{(x, i)} \left[ V \left( X \left( \sigma_D^{(n)}(t) \right), Y \left( \sigma_D^{(n)}(t) \right) \right) - V(x, i) \right] \\ &= \mathbb{E}_{(x, i)} \left[ \int_0^{\sigma_D^{(n)}(t)} \mathcal{L}V(X(s), Y(s)) ds \right] = -\mathbb{E}_{(x, i)} \left[ \sigma_D^{(n)}(t) \right]. \end{aligned}$$

Nótese que la función  $V$  es no negativa, por lo que  $\mathbb{E}_{(x,i)} [\sigma_D^{(n)}(t)] \leq V(x, i)$ . Mientras que, dado que el proceso  $Z(t) = (X(t), Y(t))$  es regular, se sigue que  $\sigma_D^{(n)}(t) \rightarrow \sigma_D(t)$  c.s. cuando  $n \rightarrow \infty$ , donde  $\sigma_D(t) = \sigma_D \wedge t$ . En virtud del lema de Fatou (véase **Lema 1.4.14**), se tiene que:

$$\mathbb{E}_{(x,i)} [\sigma_D(t)] \leq V(x, i). \quad (125)$$

El argumento que aparece después de la ecuación (91) en la prueba del **Teorema 2.4.2** implica que  $\mathbb{E}_{(x,i)} [\sigma_D(t)] \leq V(x, i) < \infty$ . Por lo que el **Lema 2.4.9** implica que  $\mathbb{E}_{(x,i)} [\sigma_U] = \mathbb{E}_{(x,i)} [\sigma_{D \times \{t\}}] < \infty$ . Dado que  $(x, i) \in D^c \times S$  es arbitrario, se concluye que  $Z(t)$  es positivo recurrente con respecto a  $U$ .

*Paso 2*

Demostrar que  $u(x, i) := \mathbb{E}_{(x,i)} [\sigma_D]$  es la solución más pequeña posible a (124). Para lograrlo, sea  $n_0$  tal y como se definió previamente. Para  $n > n_0$  sea  $\sigma_D^{(n)} = \sigma_D \wedge \beta_n$ . Claramente, tenemos que para todo  $n \geq n_0$  sucede que  $\sigma_D^{(n)} \leq \sigma_D^{(n+1)}$ . Por lo que la regularidad del proceso  $Z(t)$  implica que cuando  $n \rightarrow \infty$ ,

$$\mathbb{E}_{(x,i)} [\sigma_D^{(n)}] \rightarrow \mathbb{E}_{(x,i)} [\sigma_D]. \quad (126)$$

Nótese que el *Paso 1* implica que  $\mathbb{E}_{(x,i)} [\sigma_D] < \infty$ . Por otro lado, el **Lema 2.4.5** implica que la función  $u_n(x, i) = \mathbb{E}_{(x,i)} [\sigma_D^{(n)}]$  es una solución al problema de condiciones fronterizas:

$$\begin{cases} \mathcal{L}u_n(x, i) = -1, \\ u_n(x, i) \Big|_{x \in \partial D} = 0, \\ u_n(x, i) \Big|_{|x|=n} = 0, i \in S. \end{cases} \quad (127)$$

Por lo que la función  $v_n(x, i) := u_{n+1}(x, i) - u_n(x, i)$  es  $\mathcal{L}$ -armónica dentro del dominio  $(D^c \cap \{|x| < n\}) \times S$ . Como  $\sigma_D^{(n)} \leq \sigma_D^{(n+1)}$ , se sigue que  $\mathbb{E}_{(x,i)} [\sigma_D^{(n)}] \leq \mathbb{E}_{(x,i)} [\sigma_D^{(n+1)}]$ , por ende  $v_n(x, i) \geq 0$ . Además, (126) implica que

$$u(x, i) = u_{n_0}(x, i) + \sum_{k=n_0}^{\infty} v_k(x, i). \quad (128)$$

Ocupando la desigualdad de Harnack para sistemas de ecuaciones  $\mathcal{L}$ -elípticas (véase **Proposición 1.4.7** [33], [3] y [10]) puede demostrarse al ocupar una ligera modificación de los argumentos usuales ([19], pp. 21-22) que la suma de una serie convergente de funciones  $\mathcal{L}$ -armónicas es también una función  $\mathcal{L}$ -armónica. Con lo cual se concluye que  $u(x, i)$  es doblemente diferenciable y satisface la ecuación (124). Para verificar que  $u(x, i)$  es la solución positiva más pequeña de (124), sea  $w(x, i)$  cualquier solución positiva de (124). Nótese que  $u_n(x, i) = \mathbb{E}_{(x,i)} [\sigma_D^{(n)}]$  satisface las condiciones de frontera:

$$\begin{aligned} u_n(x, i) \Big|_{x \in \partial D} &= 0, \\ u_n(x, i) \Big|_{|x|=n} &= 0, i \in S. \end{aligned}$$

Entonces las funciones  $u_n(x, i) - w(x, i)$ ,  $i \in S$  son  $\mathcal{L}$ -armónicas y satisfacen  $w(x, i) = u_n(x, i)$  para todo  $(x, i) \in \partial D \times S$  y  $w(x, i) > u_n(x, i)$  para todo  $(x, i) \in \{|x| = n\} \times S$ . Por lo tanto se sigue del principio del máximo para sistemas de ecuaciones  $\mathcal{L}$ -elípticas [29] que  $u_n(x, i) < w(x, i)$  en  $(D^c \cap \{|x| < n\}) \times S$  para todo  $n \geq n_0$ . Haciendo  $n \rightarrow \infty$ , obtendremos que  $u(x, i) \leq w(x, i)$ , como es deseado.

*Paso 3*

Demostrar que existe una función no negativa  $V$  que satisface las condiciones del teorema si el proceso  $Z(t) = (X(t), Y(t))$  es positivo recurrente con respecto al dominio  $U = D \times \{l\}$ . Entonces  $\mathbb{E}_{(x,i)}[\sigma_D] < \infty$  para todo  $(x, i) \in D^c \times S$  y, en consecuencia, la ecuación (128) y la desigualdad de Harnack para sistemas de ecuaciones  $\mathcal{L}$ -elípticas (véase **Proposición 1.4.7**), implican que la sucesión monótona, acotada y creciente  $u_n(x, i)$  converge de manera uniforme en cualquier subconjunto compacto de  $D^c \times S$ . Además, su límite  $u(x, i)$  satisface la ecuación  $\mathcal{L}u(x, i) = -1$  para todo  $i \in S$ . Por lo tanto la función  $V(x, i) := u(x, i)$  satisface la ecuación (123). Con lo cual se concluye la prueba.

Q.E.D.

**Teorema 2.4.22:** *Una condición necesaria y suficiente para la recurrencia positiva con respecto al dominio  $U = D \times \{l\} \subset \mathbb{R}^r \times S$  es que para todo  $i \in S$ , existe una función no negativa  $V(\cdot, i) : D^c \rightarrow \mathbb{R}$  tal que  $V(\cdot, i)$  es continua y doblemente diferenciable y existe una constante  $\gamma > 0$ ,*

$$\mathcal{L}V(x, i) \leq -\gamma, \text{ para todo } (x, i) \in D^c \times S. \quad (129)$$

### Demostración

#### Necesidad

Esta parte se sigue directamente de la necesidad en el **Teorema 2.4.21** con  $\gamma = 1$ .

#### Suficiencia

Supóngase que existe una función no negativa  $V$  que satisface las condiciones del teorema. Defínase el tiempo de parada  $\sigma_D^{(n)}(t) = \sigma_D \wedge t \wedge \beta_n$  como en la prueba del **Teorema 2.4.21**. Por lo que la fórmula de Dynkin y la ecuación (129) implican que para todo  $(x, i) \in D^c \times S$  se tiene que:

$$\begin{aligned} & \mathbb{E}_{(x,i)} \left[ V \left( X \left( \sigma_D^{(n)}(t) \right), Y \left( \sigma_D^{(n)}(t) \right) \right) - V(x, i) \right] \\ &= \mathbb{E}_{(x,i)} \left[ \int_0^{\sigma_D^{(n)}(t)} \mathcal{L}V(X(s), Y(s)) ds \right] \leq -\gamma \mathbb{E}_{(x,i)} \left[ \sigma_D^{(n)}(t) \right]. \end{aligned}$$

Entonces, gracias a la no negatividad de la función  $V$ , tenemos que  $\mathbb{E}_{(x,i)} \left[ \sigma_D^{(n)}(t) \right] \leq \frac{V(x,i)}{\gamma}$ . Por otro lado, la regularidad del proceso  $Z(t) = (X(t), Y(t))$  implica que  $\sigma_D^{(n)}(t) \rightarrow \sigma_D(t)$  c.s. cuando  $n \rightarrow \infty$ , donde  $\sigma_D(t) = \sigma_D \wedge t$ . Por lo que el Lema de Fatou (véase **Lema 1.4.14**) implica que  $\mathbb{E}_{(x,i)}[\sigma_D(t)] \leq \frac{V(x,i)}{\gamma}$ . Además, de la prueba del **Teorema 2.4.21**, se tiene que  $\sigma_D(t) \rightarrow \sigma_D$  c.s. cuando  $t \rightarrow \infty$ . Por lo tanto se tiene que

$$\mathbb{E}_{(x,i)}[\sigma_D] \leq \frac{V(x,i)}{\gamma},$$

al aplicar el Lema de Fatou nuevamente. Entonces el **Lema 2.4.9** implica que  $\mathbb{E}_{(x,i)}[\sigma_U] = \mathbb{E}_{(x,i)}[\sigma_{D \times \{l\}}] < \infty$ . Dado que  $(x, i) \in D^c \times S$  es arbitrario, podemos concluir que  $Z(t)$  es positivo recurrente con respecto a  $U$ . Con lo cual se concluye la prueba de este teorema.

Q.E.D.

**Nótese** que los criterios de recurrencia positiva permiten establecer las trayectorias de los procesos. Supóngase que  $(X(t), Y(t))$  es positivo recurrente y que las funciones de Liapunov  $V(x, i)$  (donde  $i \in S$ ) están dadas por el **Teorema 2.4.20**, al igual que el conjunto  $D$ . Sea  $D_0$  un conjunto abierto, acotado y de cerradura compacta que satisface  $D \subset D_0$ , sea  $\tau_1$  del primer tiempo de salida de  $(X(t), Y(t))$  del conjunto  $D_0 \times S$ ; es decir,  $\tau_1 = \min\{t > 0 : X(t) \notin D_0\}$ . Definimos  $\sigma_1 = \min\{t > \tau_1 : X(t) \in D_0\}$ . Obtendremos lo siguiente:

$$\begin{aligned} \mathbb{P} \left[ \sup_{\tau_1 \leq t \leq \sigma_1} \{V(X(t), Y(t))\} \geq \gamma \right] &\leq \frac{\mathbb{E}[V(X(\tau_1), Y(\tau_1))]}{\gamma}, \text{ para } \gamma > 0, \\ \mathbb{E}[\sigma_1 - \tau_1] &\leq \frac{\mathbb{E}[V(X(\tau_1), Y(\tau_1))]}{\gamma}, \end{aligned} \quad (130)$$

donde  $\gamma$  está determinada en el **Teorema 2.4.20**. La idea puede continuarse de la siguiente forma para  $k \geq 1$ . Definimos:

$$\tau_{k+1} = \min\{t > \sigma_k : X(t) \notin D_0\},$$

y

$$\sigma_{k+1} = \min\{t > \tau_{k+1} : X(t) \in D_0\}.$$

Entonces podemos obtener las estimaciones de  $\mathbb{E}[\sigma_{k+1} - \tau_{k+1}]$ , que nos da la diferencia esperada del  $(k+1)$ -ésimo tiempo de retorno y tiempo de salida.

Ahora, con el siguiente teorema procederemos a estudiar la recurrencia positiva de sistemas linealizables (en el componente continuo) tal y como se describe en la condición **(A1.5.6)**.

**Teorema 2.4.23:** *Suponga que se satisfacen **(A1.5.5)** y **(A1.5.6)**. Si*

$$\sum_{i=1}^N \pi_i \lambda_{\max} \left( b(i) + b^*(i) + \sum_{j=1}^d \sigma_j(i) \sigma_j^*(i) \right) < 0, \quad (131)$$

entonces el proceso  $(X(t), Y(t))$  es positivo recurrente.

**Demostración:**

Para facilitar la notación, defínase el vector columna

$$\mu = [\mu_1, \mu_2, \dots, \mu_N]^* \in \mathbb{R}^N,$$

donde

$$\mu_i = \frac{1}{2} \lambda_{\max} \left( b(i) + b^*(i) + \sum_{j=1}^d \sigma_j(i) \sigma_j^*(i) \right).$$

Sea

$$\beta := -\pi \mu = -\frac{1}{2} \sum_{i=1}^N \pi_i \lambda_{\max} \left( b(i) + b^*(i) + \sum_{j=1}^d \sigma_j(i) \sigma_j^*(i) \right).$$

Nótese que (131) implica que  $\beta > 0$ . Dado que

$$\pi(\mu + \beta \mathbf{1}) = \pi \mu + \beta \cdot \pi \mathbf{1} = -\beta + \beta = 0,$$

donde  $\mathbf{1} = [1, 1, \dots, 1]^t$ , la condición **(A1.5.6)** implica que la ecuación:

$$\hat{Q}c = \mu + \beta \mathbf{1},$$

tiene una solución  $c = [c_1, c_2, \dots, c_N]^* \in \mathbb{R}^N$ , donde  $\hat{Q} = [\hat{q}_{ij}]$  es el generador de la cadena de Markov a tiempo continuo  $\hat{Y}(t)$  que se menciona en **(A1.5.6)**. Por lo tanto:

$$\mu_i - \sum_{j=1}^N \hat{q}_{ij} c_j = -\beta, i \in S. \quad (132)$$

Para cada  $i \in S$ , considérese la función de Liapunov

$$V(x, i) = (1 - \gamma c_i) |x|^\gamma,$$

donde  $0 < \gamma < 1$  es lo suficientemente pequeño tal que  $1 - \gamma c_i > 0$  para cada  $i \in S$ . Se sigue de inmediato que para cada  $i \in S$ ,  $V(\cdot, i)$  es continua, no negativa y posee segundas derivadas parciales con respecto a  $x$  en cualquier vecindad hueca de 0 (es decir, una vecindad de 0 que no contiene a 0). Haciendo los cálculos precisos, se tiene que para todo  $x \neq 0$

$$\begin{aligned} \nabla V(x, i) &= (1 - \gamma c_i) \gamma |x|^{\gamma-2} x, \\ \nabla^2 V(x, i) &= (1 - \gamma c_i) \gamma [|x|^{\gamma-2} I + (\gamma - 2) |x|^{\gamma-4} x x^*]. \end{aligned}$$

Por otro lado, se sigue de (118) que

$$\frac{a(x, i)}{|x|^2} = \frac{\sigma(x, i) \sigma^*(x, i)}{|x|^2} = \sum_{j=1}^d \left[ \frac{\sigma_j(i) x x^* \sigma_j^*(i)}{|x|^2} \right] + o(1),$$

de forma que  $o(1) \rightarrow 0$  cuando  $|x| \rightarrow \infty$ . Por lo tanto se tiene que

$$\begin{aligned} \mathcal{L}V(x, i) &= \frac{\gamma}{2} (1 - \gamma c_i) \sum_{j=1}^d \left( x^* \sigma_j^*(i) |x|^{\gamma-2} I \sigma_j(i) x + x^* \sigma_j^*(i) (\gamma - 2) |x|^{\gamma-4} x x^* \sigma_j(i) x \right) \\ &\quad + (1 - \gamma c_i) \gamma |x|^{\gamma-2} x^* b(i) x + o(|x|^\gamma) - \sum_{i \neq j} q_{ij}(x) |x|^\gamma (c_j - c_i) \\ &= \gamma (1 - \gamma c_i) |x|^\gamma \left\{ \frac{1}{2} \sum_{j=1}^d \left( \frac{x^* \sigma_j^*(x, i) \sigma_j(x, i) x}{|x|^2} + (\gamma - 2) \frac{(x^* \sigma_j^*(i) x)^2}{|x|^4} \right) + \frac{x^* b(i) x}{|x|^2} - \sum_{i \neq j} q_{ij}(x) \frac{c_j - c_i}{1 - \gamma c_i} + o(1) \right\}, \end{aligned} \quad (133)$$

donde  $o(1) \rightarrow 0$  cuando  $|x| \rightarrow \infty$ . Nótese que

$$\frac{x^* b(i) x}{|x|^2} + \frac{1}{2} \sum_{j=1}^d \left[ \frac{x^* \sigma^*(x, i) \sigma(x, i) x}{|x|^2} \right] \leq \frac{1}{2} \lambda_{\max} \left( b(i) + b^*(i) \sum_{j=1}^d \sigma_j^*(i) \sigma_j(i) \right) = \mu_i. \quad (134)$$

Ahora, ocupando **(A1.5.6)**

$$\sum_{j \neq i} q_{ij}(x) \frac{c_j - c_i}{1 - \gamma c_i} = \sum_{j=1}^N q_{ij}(x) c_j + \sum_{j \neq i} q_{ij}(x) \frac{c_i (c_j - c_i)}{1 - \gamma c_i} \gamma = \sum_{j=1}^N \hat{q}_{ij} c_j + O(\gamma) + o(1), \quad (135)$$

donde  $\frac{O(\gamma)}{\gamma}$  está acotado y  $o(1) \rightarrow 1$  cuando  $|x| \rightarrow \infty$  y  $\gamma \rightarrow 0$ . Por lo tanto se sigue de (133)-(135) que cuando  $|x| > R$  donde  $R$  es lo suficientemente grande y  $0 < \gamma < 1$  lo suficientemente pequeño, se tiene que

$$\mathcal{L}X(x, i) \leq \gamma(1 - \gamma c_i)|x|^\gamma \left[ \mu_i - \sum_{j=1}^N \hat{q}_{ij} c_j + O(\gamma) + o(1) \right].$$

Además, en virtud de (132), se tiene que

$$\mathcal{L}V(x, i) \leq \gamma(1 - \gamma c_i)|x|^\gamma (-\beta + O(\gamma) + o(1)) \leq -K < 0,$$

para todo  $(x, i) \in \mathbb{R}^r \times S$  donde  $|x| > r$ , donde  $K$  es una constante positiva. Por lo tanto se concluye del **Teorema 2.4.22** que el proceso  $(X(t), Y(t))$  es positivo recurrente.

Q.E.D.

El resultado anterior posee el siguiente caso particular (véase [36]).

**Corolario 2.4.24:** *Supóngase que el componente continuo  $X(t)$  es unidimensional y que si  $|x| \rightarrow \infty$ ,*

$$\begin{aligned} \frac{b(x, i)}{|x|} &= b_i \frac{x}{|x|} + o(1), \\ \frac{\sigma(x, i)}{|x|} &= \sigma_i \frac{x}{|x|} + o(1). \end{aligned} \tag{136}$$

para algunas constantes  $b_i, \sigma_i, i \in S$ . Si

$$\pi b - \frac{1}{2} \pi \sigma^2 := \sum_{i=1}^N \pi_i \left( b_i - \frac{\sigma_i^2}{2} \right) < 0, \tag{137}$$

de forma que  $\pi = [\pi_1, \dots, \pi_N] \in \mathbb{R}^{1 \times N}$  es como en la condición **(A1.5.6)**, entonces de proceso  $(X(t), Y(t))$  es positivo recurrente.

La siguiente es una condición suficiente para la recurrencia no positiva.

**Teorema 2.4.25:** *Supóngase que existe un dominio no vacío y acotado  $D$  tal que existen funciones  $V(x, i)$  y  $W(x, i)$  definidas en  $D^c \times S$  que satisfacen:*

- (a)  $V(x, i) \geq 0$  para todo  $(x, i) \in D^c \times S$ , y para alguna constante positiva  $k$ ,  
 $0 \leq \mathcal{L}V(x, i) \leq k$  para todo  $(x, i) \in D^c \times S$ ;
- (b)  $W(x, i) \leq 0$  para todo  $(x, i) \in \partial D \times S$ , y  
 $\mathcal{L}W(x, i) \geq 0$  para todo  $(x, i) \in D^c \times S$ ;
- (c) Para alguna sucesión creciente de dominios acotados  $D \subset E_n$  con bordes  $\Gamma_n$ ,  
 $\frac{\inf_{(x, i) \in \Gamma_n \times S} \{V(x, i)\}}{\sup_{(x, i) \in \Gamma_n \times S} \{W(x, i)\}} = R_n \rightarrow \infty$  si  $n \rightarrow \infty$ .

Si existe algún  $(x, \alpha) \in D^c \times S$  que satisface  $W(x, \alpha) > 0$ , entonces el proceso  $(X(t), Y(t))$  no es positivo recurrente. Eso es, que  $\mathbb{E}[\sigma_D] = \infty$  para todo  $(x, \alpha) \in D^c \times S$  tal que  $W(x, \alpha) > 0$ , donde:

$$\sigma_D = \sigma_D^{x, \alpha} = \inf\{t \geq 0 : X^{x, \alpha}(t) \in D\}.$$

**Demostración:**

Considérese la función  $V - R_n W$  definida en  $(D^c \cap E_n) \times S$ . Entonces se sigue de las condiciones (a), (b) y (c) que



$$\begin{cases} \mathcal{L}(V - R_n W)(x, i) \leq k, \\ (V - R_n W)(x, i) \Big|_{x \in \partial D} \geq 0, \\ (V - R_n W)(x, i) \Big|_{x \in \Gamma_n} \geq 0. \end{cases}$$

Fijando cualquier  $(x, \alpha) \in D^c \times S$  que satisfice  $W(x, \alpha) > 0$ , entonces se sigue que  $\mathbb{P}[\sigma_D \wedge \beta_n < \infty] = 1$  para todo  $n > (n_0 \vee |y|)$ , donde  $n_0$  es un entero lo suficientemente grande tal que  $D \subset \{x \in \mathbb{R}^r : |x| < n_0\}$  y  $\beta_n$  tal y como se define en (44).

Por lo que la fórmula de Dynkin implica que

$$\mathbb{E}[(V - R_n W)(X(\sigma_D \wedge \beta_n), Y(\sigma_D \wedge \beta_n)) - (V - R_n W)(x, \alpha)] \leq k \mathbb{E}[\sigma_D \wedge \beta_n].$$

Es decir,

$$\mathbb{E}[\sigma_D \wedge \beta_n] \geq \frac{R_n W(x, \alpha) - V(x, \alpha)}{k}.$$

Por lo tanto, se sigue de la condición (c) que  $\mathbb{E}[\sigma_D \wedge \beta_n] \rightarrow \infty$ , cuando  $n \rightarrow \infty$ . Por otro lado, nótese que  $\sigma_D \geq \sigma_D \wedge \beta_n$ . Se tiene que  $\mathbb{E}[\sigma_D] = \infty$ . Por lo tanto, el proceso  $(X(t), Y(t))$  no es positivo recurrente en virtud del **Teorema 2.4.10**.

Q.E.D.

Por otro lado, la recurrencia nula es más complicada. Para simplificar el asunto, vamos a suponer que el componente continuo del proceso se compone de una difusión sin coeficiente de deriva.

**Teorema 2.4.26:** *Considérese un proceso de difusión cambiante de variable real, eso es:*

$$\begin{cases} dX(t) = \sigma(X(t), Y(t))dw(t), \\ \mathbb{P}[Y(t + \Delta) = j | Y(t) = i, X(s), Y(s), s \leq t] = q_{ij}(X(t))\Delta + o(\Delta). \end{cases} \quad (138)$$

Si para algunas constantes  $0 \leq \beta \leq 1, k_1 > 0$ , y  $k_2 > 0$ ,

$$\sigma^2(x, i) \leq k_1 |x|^{1-\beta}, \text{ para todo } (x, i) \in \{x \in \mathbb{R}, |x| \geq k_2\} \times S, \quad (139)$$

entonces el proceso  $(X(t), Y(t))$  definido en (138) es recurrente nulo.

**Demostración:**

De hecho, gracias al **Teorema 2.4.10**, basta con probar que el proceso es recurrente nulo con respecto a  $D \times \{l\}$ , donde  $D = (-k_2, k_2)$  y  $l \in S$ . Para todo  $i \in S$ , la función de Liapunov

$$V(x, i) = \begin{cases} |x|^{1+\beta}, & \text{si } \beta > 0, \\ |x|(\ln(|x|) - 1) + k, & \text{si } \beta = 0, \end{cases}$$

donde  $k$  es una constante positiva lo suficientemente grande, y

$$W(x, i) = |x| - k_2 \text{ para } (x, i) \in D^c \times S.$$

Los cálculos exactos muestran que la función  $W$  satisfice (101) en el **Teorema 2.4.11**. Esto implica que el proceso es recurrente con respecto a  $D \times \{l\}$ .

De forma similar, en virtud de (139), los cálculos exactos muestran que las funciones  $V$  y  $W$  satisfacen las condiciones (a)-(c) del **Teorema 2.4.25**. Por lo tanto, el proceso definido en (138) es recurrente nulo.

Q.E.D.

**Nótese** que si para algunas constantes positivas  $\rho, k_1$ , y  $k_2$ , tenemos que:

$$\sigma^2(x, i) \leq k_1 |x|^{1+\rho}, \text{ para todo } (x, i) \in \{x \in \mathbb{R}, |x| \geq k_2\} \times S, \quad (140)$$

entonces el proceso  $(X(t), Y(t))$  definido en (138) es positivo recurrente.

Para verificarlo, seleccionamos  $1 > \delta > 0$  de forma tal que  $\delta + \rho > 1$  y definimos  $V(x, i) = |x|^\delta$  para todo  $(x, i) \in \{x \in \mathbb{R}, |x| \geq k_2\} \times S$ . Entonces podemos probar que

$$\mathcal{L}V(x, i) = \frac{1}{2} \delta(\delta - 1) |x|^{\delta-2} \sigma^2(x, i) \leq \frac{1}{2} k_1 \delta(\delta - 1) k_2^{\delta+\rho-1} < 0,$$

donde  $(x, i) \in \{x \in \mathbb{R} : |x| \geq k_2\} \times S$ . Por lo tanto, se sigue del **Teorema 2.4.22** que el proceso  $(X(t), Y(t))$  es positivo recurrente.

Para finalizar esta sección se presentará una serie de simulaciones programadas en R para evidenciar algunos patrones de comportamiento de los procesos de difusión cambiante, en este caso el análisis de recurrencia.

**Ejemplo 2.4.27:** *Considérese el proceso  $(X(t), Y(t))$  que satisface las siguientes condiciones:*

$$\begin{cases} dX(t) = b(X(t), Y(t))dt + \sigma(X(t), Y(t))dw(t), \\ \mathbb{P}[Y(t + \Delta) = j | Y(t) = i, X(s), Y(s), s \leq t, t \geq 0] = q_{ij}(X(t))\Delta + o(\Delta) : (X(t), Y(t)) \in \mathbb{R} \times \{1, 2, 3\}, \end{cases} \quad (141)$$

donde el componente discreto  $Y(t)$  salta entre tres fases de acuerdo con el siguiente generador

$$Q(x) = \begin{bmatrix} -\frac{2+|x|}{1+x^2} - \frac{1+3x^2}{2+x^2} & \frac{2+|x|}{1+x^2} & \frac{1+3x^2}{2+x^2} \\ 1 - \frac{\sin(x)}{1+x^2} & \frac{\sin(x)}{1+x^2} - \frac{\cos(x)}{2+|x|} - 3 & 2 + \frac{\cos(x)}{2+|x|} \\ \frac{1+\cos^2(x)}{2+x^2} & 2 & -\frac{1+\cos^2(x)}{2+x^2} - 2 \end{bmatrix},$$

y los coeficientes de deriva y de difusión son los siguientes

$$\begin{aligned} b(x, 1) &= 3x - 1, b(x, 2) = \frac{3}{2}x + 1, b(x, 3) = -x + \frac{x}{1+|x|} \\ \sigma(x, 1) &= \sqrt{3+x^2}, \sigma(x, 2) = \sqrt{2 - \sin(x) + 2x^2}, \sigma(x, 3) = 3 + \sqrt{4+x^2}. \end{aligned}$$

Por lo que, de acuerdo a (141), hay tres procesos de difusión

$$dX(t) = (3X(t) - 1)dt + \sqrt{3 + X(t)^2}dw(t) \quad (142)$$

$$dX(t) = \left(\frac{3}{2}X(t) + 1\right)dt + \sqrt{2 - \sin(X(t)) + 2X(t)^2}dw(t) \quad (143)$$

$$dX(t) = \left(-X(t) + \frac{X(t)}{1+|X(t)|}\right)dt + \left(3 + \sqrt{4 + X(t)^2}\right)dw(t), \quad (144)$$

y la dinámica de  $(X(t), Y(t))$  implica saltar de un proceso a otro de acuerdo a los saltos de la componente  $Y(t)$ .

En virtud del Teorema 3.7.3 de [27], podemos garantizar que (142) es positivo recurrente, mientras que (143) y (144) son transitivos. Pero, debido al efecto de estabilización del proceso  $Y(t)$ , el proceso (141) es positivo recurrente. Procederemos a demostrar, ocupando el **Corolario 2.4.24**, que efectivamente (141) es positivo recurrente. De hecho, puede observarse que si  $|x| \rightarrow \infty$ , las constantes  $b(i)$  y  $\sigma^2(i)$ ,  $i = 1, 2, 3$ , en (136) están dadas por

$$\begin{aligned} b(1) = 3, b(2) = \frac{3}{2}, b(3) = -1, \\ \sigma^2(1) = 1, \sigma^2(2) = 3, \sigma^2(3) = 1. \end{aligned} \tag{145}$$

Además, si  $|x| \rightarrow \infty$ , entonces  $Q(x)$  tenderá a

$$\hat{Q} = \begin{bmatrix} -3 & 0 & 3 \\ 1 & -3 & 2 \\ 0 & 2 & -2 \end{bmatrix}.$$

Al resolver el sistema de ecuaciones

$$\begin{cases} \pi \hat{Q} = 0, \\ \pi \cdot [1, 1, 1]^t = 1, \end{cases}$$

obtendremos que la distribución estacionaria  $\pi$  asociada con  $\hat{Q}$  es

$$\pi = \left( \frac{2}{17}, \frac{6}{17}, \frac{9}{17} \right). \tag{146}$$

Por lo que en virtud de (145) y (146), se puede observar que

$$\sum_{i=1}^3 \pi_i \left( b(i) - \frac{\sigma^2(i)}{2} \right) = -\frac{23}{34} < 0.$$

Por lo que el **Corolario 2.4.24** implica que (141) es positivo recurrente; véase la trayectoria de prueba en la siguiente figura.

## Ejemplo de Recurrencia

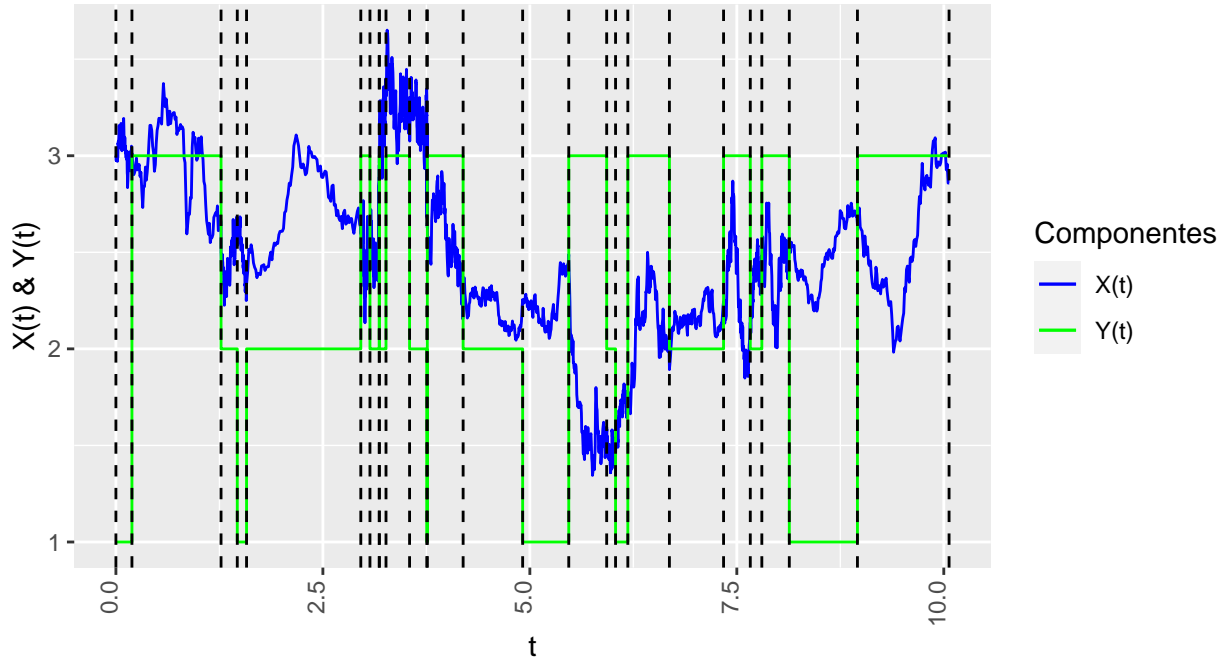


Figura 5:  
Traectoria de un proceso de Markov bivariado de tres fases y con condiciones iniciales  $(x,y)=(3,0)$ .

## 2.5 Ergodicidad

Empezaremos estudiando las propiedades de ergodicidad del proceso  $(X(t), Y(t))$  que satisface (28) y (29), suponiendo que el proceso es positivo recurrente con respecto a algún dominio acotado  $U = E \times \{l\}$ , donde  $E \subset \mathbb{R}^r$  y  $l \in S$  están fijos por lo que resta de la sección. También suponemos que la frontera  $\partial E$  de  $E$  es lo suficientemente suave y que el operador  $\mathcal{L}$  satisface **(A1.5.5)**. Entonces se sigue del **Teorema 2.4.10** que el proceso es positivo recurrente con respecto a cualquier conjunto abierto no vacío.

Sea  $D \subset \mathbb{R}^r$  un conjunto abierto con frontera suficientemente suave  $\partial D$  de forma que  $E \cup \partial E \subset D$ . Sea  $\zeta_0 = 0$ , definimos los tiempos de paro  $\zeta_1, \zeta_2, \dots$  de manera inductiva como:  $\zeta_{2n+1}$  es el primer tiempo después de  $\zeta_{2n}$  en el que el proceso  $(X(t), Y(t))$  llega al conjunto  $\partial E \times \{l\}$  y  $\zeta_{2n+2}$  es el primer tiempo después de  $\zeta_{2n+1}$  en el cual el proceso  $(X(t), Y(t))$  llega al conjunto  $\partial D \times \{l\}$ . Por lo que podemos dividir un camino arbitrario del proceso  $(X(t), Y(t))$  en ciclos:

$$[\zeta_0, \zeta_2), [\zeta_2, \zeta_4), \dots, [\zeta_{2n}, \zeta_{2n+2}), \dots \quad (147)$$

El proceso  $(X(t), Y(t))$  es positivo recurrente con respecto a  $E \times \{l\}$  y, por el **Teorema 2.4.10**, es positivo recurrente con respecto a  $D \times \{l\}$ . Se sigue que todos los tiempos de paro  $\zeta_0 < \zeta_1 < \zeta_2 < \zeta_3 < \dots$  son finitos casi seguramente (c.s.). Dado que el proceso  $(X(t), Y(t))$  es positivo recurrente, asumiremos sin pérdida de generalidad que  $(X(0), Y(0)) = (x, l) \in \partial D \times \{l\}$ . Ocupando la Propiedad Fuerte de Markov que satisface  $(X(t), Y(t))$ , tendremos que la sucesión  $\{Z_n\}$  es una cadena de Markov en  $\partial D \times \{l\}$ , donde  $Z_n = (X(\zeta_{2n}), Y(\zeta_{2n})) = (X_n, l), n = 0, 1, \dots$ . Sea  $\hat{\mathbb{P}}(x, A)$  la probabilidad de transición a un paso de esta cadena de Markov, es decir:

$$\hat{\mathbb{P}}(x, A) = \mathbb{P}[Z_1 \in A \times \{l\} | Z_0 = (x, l)],$$

para todo  $x \in \partial D$  y  $A \in \mathbb{B}(\partial D)$ , donde  $\mathbb{B}(\partial D)$  denota la colección de conjuntos Borel medibles en  $\partial D$ . Nótese que el proceso  $(X(t), Y(t))$  que comienza en  $(x, l)$ , puede saltar varias veces antes de llegar al conjunto  $(A, l)$ . Ahora, sea  $\hat{\mathbb{P}}^{(n)}(x, A)$  la probabilidad de transición a  $n$  pasos de esta cadena de Markov para  $n \geq 1$ . Para toda función Borel medible  $f : \mathbb{R}^r \rightarrow \mathbb{R}$ , definimos:

$$\mathbb{E}_x[f(X_1)] := \mathbb{E}_{(x,l)}[f(X_1)] = \int_{\partial D} f(y) \hat{\mathbb{P}}(x, dy). \quad (148)$$

Más adelante, vamos a probar que el proceso  $(X(t), Y(t))$  posee una distribución estacionaria única. Para lograrlo, primero necesitaremos los siguientes resultados:

**Lema 2.5.1:** *La cadena de Markov  $Z_i = (X_i, l)$  posee una distribución estacionaria única  $m(\cdot)$  tal que la matriz de transición de  $n$  pasos  $\hat{\mathbb{P}}^{(n)}(x, A)$  satisface que:*

$$\left| \hat{\mathbb{P}}^{(n)}(x, A) - m(A) \right| < \lambda^n, \text{ para todo } A \in \mathbb{B}(\partial D), \quad (149)$$

para alguna constante  $0 < \lambda < 1$ .

**Demostración:**

Nótese que

$$\hat{\mathbb{P}}(x, A) = \mathbb{P}[Z_1 \in A \times \{l\} | Z_0 = (x, l)] = \int_{\partial D} \mathbb{P}_{(x,l)}[(X(\zeta_1), Y(\zeta_1)) \in dy \times \{l\}] \times \mathbb{P}_{(y,l)}[(X(\zeta_2), Y(\zeta_2)) \in A \times \{l\}].$$

Ocupando la **Definición 2.3.14** de función armónica, (de manera similar a la solución de un proceso de difusión no cambiante), podemos concluir la prueba de este lema de forma análoga al Lema 4.4.1 de [27].

Q.E.D.

Nótese que:

$$(X^{s, X(s), Y(s)}(t), Y^{s, X(s), Y(s)}(t)) = (X^{0, X(0), Y(0)}(t+s), Y^{0, X(0), Y(0)}(t+s)), \quad (150)$$

donde  $(X^{0, X(0), Y(0)}(u), Y^{0, X(0), Y(0)}(u))$  denota el camino muestral del proceso  $(X(\cdot), Y(\cdot))$  que comienza en  $(X(0), Y(0))$ , tal que  $(X^{s, X(s), Y(s)}(t), Y^{s, X(s), Y(s)}(t))$  se define de manera similar. De forma que no haya confusiones, escribiremos:

$$(X(u), Y(u)) = (X^{0, X(0), Y(0)}(u), Y^{0, X(0), Y(0)}(u)).$$

Sean  $\tau$  un tiempo de parada con  $\mathbb{E}_{(x,i)}[\tau] < \infty$  y  $f : \mathbb{R}^r \times S \rightarrow \mathbb{R}$  una función Borel medible. Entonces:

$$\mathbb{E}_{(x,i)} \left[ \int_0^\tau f(X(s+t), Y(s+t)) ds \right] = \mathbb{E}_{(x,i)} \left[ \int_0^\tau \mathbb{E}_{(X(s), Y(s))} [f(X(s+t), Y(s+t))] ds \right]. \quad (151)$$

Ahora ya podemos construir una distribución estacionaria para el proceso  $Z(t) = (X(t), Y(t))$  de manera explícita.

**Teorema 2.5.2:** *El proceso positivo recurrente  $Z(t) = (X(t), Y(t))$  posee una distribución estacionaria única  $\nu(\cdot, \cdot) = (\nu(\cdot, i) : i \in S)$ .*

**Demostración:**

Recuérdese cómo se definieron los ciclos en (147). Sean  $A \in \mathbb{B}(\mathbb{R}^r)$  e  $i \in S$ . Denotaremos por  $\tau^{A \times \{i\}}$  al tiempo que pasa el camino de  $Z(t) = (X(t), Y(t))$  en el conjunto  $A \times \{i\}$  durante el primer ciclo. Sea

$$\hat{\nu}(A, i) := \int_{\partial D} m(dx) \mathbb{E} [\tau^{A \times \{i\}}], \quad (152)$$

donde  $m(\cdot)$  es la distribución estacionaria de la cadena de Markov  $Z_i = (X_i, l)$ , cuya existencia está garantizada por el **Lema 2.5.1**. Es fácil verificar que  $\hat{\nu}(\cdot, \cdot)$  es una medida positiva definida sobre  $\mathbb{B}(\mathbb{R}^r) \times S$ . Por lo que para cualquier función Borel medible y acotada  $g(\cdot) : \mathbb{R}^r \mapsto \mathbb{R}$ , se sigue de (148) y del teorema de Fubini que

$$\int_{\partial D} \mathbb{E}_x [g(X_1)] m(dx) = \int_{\partial D} \int_{\partial D} m(dx) g(y) \hat{\mathbb{P}}(x, dy) = \int_{\partial D} g(y) m(dy). \quad (153)$$

Ahora, afirmamos que para toda función acotada y continua  $f(\cdot, \cdot)$  se satisface que

$$\sum_{j=1}^N \int_{\mathbb{R}^r} f(y, j) \hat{\nu}(dy, j) = \int_{\partial D} m(dx) \mathbb{E}_x \left[ \int_0^{\zeta_2} f(X(t), Y(t)) dt \right]. \quad (154)$$

De hecho, si  $f$  es una función indicadora, es decir  $f(y, j) = \mathbb{1}_{A \times \{i\}}(y, j)$  para algún  $A \in \mathbb{B}(\mathbb{R}^r)$  e  $i \in S$ , entonces por (152) tenemos

$$\begin{aligned} \sum_{j=1}^N \int_{\mathbb{R}^r} \mathbb{1}_{A \times \{i\}}(y, j) \hat{\nu}(dy, j) &= \hat{\nu}(A, i) = \int_{\partial D} m(dx) \mathbb{E}_x [\tau^{A \times \{i\}}] \\ &= \int_{\partial D} m(dx) \mathbb{E}_x \left[ \int_0^{\zeta_2} \mathbb{1}_{A \times \{i\}}(X(t), Y(t)) dt \right]. \end{aligned}$$

De manera similar, se satisface (154) para  $f$  función simple, es decir

$$f(y, j) = \sum_{p=1}^m c_p \mathbb{1}_{U_p}(y, j),$$

donde  $U_p \subset \mathbb{R}^r \times S$ . Finalmente, si  $f$  es una función continua y acotada se satisface (154) al aproximar  $f$  por medio de funciones simples. Se sigue de las ecuaciones (154), (150) y (151) que

$$\begin{aligned} \sum_{i=1}^N \int_{\mathbb{R}^r} \mathbb{E}_{(x, i)} [f(X(t), Y(t)) \hat{\nu}(dx, i)] &= \int_{\partial D} m(dx) \mathbb{E}_x \left[ \int_0^{\zeta_2} f(X(u), Y(u)) du \right] \\ &+ \int_{\partial D} m(dx) \mathbb{E}_x \left[ \int_0^{\zeta_2+t} f(X(u), Y(u)) du \right] - \int_{\partial D} m(dx) \mathbb{E}_x \left[ \int_0^t f(X(u), Y(u)) du \right]. \end{aligned}$$

Ahora, aplicando (153) con

$$g(x) = \mathbb{E}_x \left[ \int_{\zeta_2}^{\zeta_2+t} f(X(u), Y(u)) du \right],$$

tendremos que

$$\int_{\partial D} m(dx) \mathbb{E}_x \left[ \int_{\zeta_2}^{\zeta_2+t} f(X(u), Y(u)) du \right] = \int_{\partial D} m(dx) \mathbb{E}_x \left[ \int_0^t f(X(u), Y(u)) du \right].$$

Nótese que en la anterior deducción se ocupó (150) nuevamente. Por lo tanto, las dos ecuaciones anteriores y (154) implican que:

$$\sum_{i=1}^N \int_{\mathbb{R}^r} \mathbb{E}_{(x,i)} [f(X(t), Y(t)) \hat{\nu}(dx, i)] = \sum_{i=1}^N \int_{\mathbb{R}^r} f(x, i) \hat{\nu}(dx, i).$$

Por lo que la medida normalizada

$$\nu(A, i) = \frac{\hat{\nu}(A, i)}{\sum_{j=1}^N \hat{\nu}(\mathbb{R}^r, j)}, \quad i \in S, \quad (155)$$

define la distribución estacionaria deseada. Con lo cual se concluye la prueba.

Q.E.D.

**Teorema 2.5.3:** Denotemos por  $\mu(\cdot, \cdot)$  a la densidad estacionaria asociada a la distribución estacionaria  $\nu(\cdot, \cdot)$  construida en el **Teorema 2.5.2** y sea  $f(\cdot, \cdot) : \mathbb{R}^r \times S \rightarrow \mathbb{R}$  una función Borel medible tal que:

$$\sum_{i=1}^N \int_{\mathbb{R}^r} |f(x, i)| \mu(x, i) dx < \infty. \quad (156)$$

Entonces:

$$\mathbb{P}_{(x,i)} \left( \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X(t), Y(t)) dt = \bar{f} \right) = 1, \quad (157)$$

donde para cualquier  $(x, i) \in \mathbb{R}^r \times S$ :

$$\bar{f} = \sum_{i=1}^N \int_{\mathbb{R}^r} f(x, i) \mu(x, i) dx. \quad (158)$$

**Demostración:**

Primero demostraremos (157) para el caso donde la distribución inicial es la distribución estacionaria de la cadena de Markov  $Z_i = (X_i, Y_i)$ , es decir:

$$\mathbb{P}[(X(0), Y(0)) \in A \times \{l\}] = m(A), \quad (159)$$

para todo  $A \in \mathbb{B}(\partial D)$ . Considérese la siguiente sucesión de variables aleatorias

$$\eta_n = \int_{\zeta_{2n}}^{\zeta_{2n+2}} f(X(t), Y(t)) dt. \quad (160)$$

Por lo que se sigue de la ecuación (159) que  $\{\eta_n\}$  es una sucesión estacionaria. Además, se sigue de (152) y (154) que

$$\mathbb{E}[\eta_n] = \sum_{i=1}^N \int_{\mathbb{R}^r} f(x, i) \hat{\nu}(dx, i), \quad (161)$$

para todo  $n = 0, 1, 2, \dots$ . Por otro lado, la ecuación (149) implica que la sucesión  $\{\eta_n\}$  es transitiva. Sea  $\nu(T)$  el número de ciclos completados hasta el tiempo  $T$ . Es decir:

$$\nu(T) := \max \left\{ n \in \mathbb{N} : \sum_{k=1}^n (\zeta_{2k} - \zeta_{2k-2}) \leq T \right\}.$$

Por lo que puede descomponerse  $\int_0^T f(X(t), Y(t)) dt$  en

$$\int_0^T f(X(t), Y(t)) dt = \sum_{n=0}^{\nu(T)} \eta_n + \int_{\zeta_{2\nu(T)}}^T f(X(t), Y(t)) dt, \quad (162)$$

donde  $\eta_n$  está dado por (160). Podemos asumir sin pérdida de generalidad que  $f(x, i) \geq 0$  (para el caso general, podemos expresar  $f(x, i)$  como la diferencia entre dos funciones no negativas). Por lo que se sigue de (162) que

$$\sum_{n=0}^{\nu(T)} \eta_n \leq \int_0^T f(X(t), Y(t)) dt \leq \sum_{n=0}^{\nu(T)+1} \eta_n.$$

En vista de que la sucesión  $\{\eta_n\}$  es estacionaria y transitiva, la ley de los grandes números (para este caso de sucesiones) implica que

$$\mathbb{P} \left[ \frac{1}{n} \sum_{k=0}^n \eta_k \rightarrow \sum_{i=1}^N \int_{\mathbb{R}^r} f(x, i) \hat{\nu}(dx, i), \text{ cuando } n \rightarrow \infty \right] = 1. \quad (163)$$

En particular, si  $f(x, i) = 1$ , entonces la anterior ecuación se reduce a

$$\mathbb{P} \left[ \frac{\zeta_{2n+2}}{n} \rightarrow \sum_{i=1}^N \hat{\nu}(\mathbb{R}^r, i), \text{ cuando } n \rightarrow \infty \right] = 1. \quad (164)$$

Nótese que la recurrencia positiva del proceso  $Z(t) = (X(t), Y(t))$  implica que  $\nu(T) \rightarrow \infty$  cuando  $T \rightarrow \infty$ . Claramente  $\frac{\nu(T)}{\nu(T)+1} \rightarrow 1$  casi seguramente cuando  $T \rightarrow \infty$ . Por lo que, se sigue de (164), cuando  $T \rightarrow \infty$

$$\frac{\zeta_{2\nu(T)}}{\zeta_{2\nu(T)+2}} = \frac{\frac{\zeta_{2\nu(T)}}{\nu(T)}}{\frac{\zeta_{2\nu(T)+2}}{\nu(T)+1}} \frac{\nu(T)}{\nu(T)+1} \rightarrow 1 \text{ c.s.} \quad (165)$$

Por otro lado, dado que  $\zeta_{2\nu(T)} \leq T \leq \zeta_{2\nu(T)+2}$ , se tiene que

$$\frac{\zeta_{2\nu(T)}}{\zeta_{2\nu(T)+2}} \leq \frac{\zeta_{2\nu(T)}}{T} \leq \frac{\zeta_{2\nu(T)}}{\zeta_{2\nu(T)}} = 1.$$

Por ende, se tiene de (165) que

$$\frac{\zeta_{2\nu(T)}}{T} \rightarrow 1 \text{ c.s. cuando } T \rightarrow \infty. \quad (166)$$



Además, (164) implica que

$$\frac{\nu(T)}{\zeta_{2\nu(T)}} \rightarrow \frac{1}{\sum_{i=1}^N \hat{\nu}(\mathbb{R}^r, i)} \text{ c.s cuando } T \rightarrow \infty. \quad (167)$$

Por último, nótese que

$$\int_{\mathbb{R}^r} f(x, i) \nu(dx, i) = \int_{\mathbb{R}^r} f(x, i) \mu(x, i) dx,$$

por la definición de  $\mu(\cdot, \cdot)$ . Por lo que se satisface la ecuación (157), con lo cual queda demostrado (157) en el caso donde la distribución inicial es (159).

Sea  $(x, i) \in \mathbb{R}^r \times S$ . Dado que el proceso  $Z(t) = (X(t), Y(t))$  es positivo recurrente con respecto al dominio  $D \times \{l\}$ , se tiene que

$$\mathbb{P}_{(x,i)} \left[ \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X(t), Y(t)) dt = \bar{f} \right] = \mathbb{P}_{(x,i)} \left[ \lim_{T \rightarrow \infty} \frac{1}{T} \int_{\zeta_2}^T f(X(t), Y(t)) dt = \bar{f} \right].$$

Podemos reescribir la expresión anterior como

$$\begin{aligned} \mathbb{P}_{(x,i)} \left[ \lim_{T \rightarrow \infty} \frac{1}{T} \int_{\zeta_2}^T f(X(t), Y(t)) dt = \bar{f} \right] &= \\ \int_{\partial D} \mathbb{P}_{(x,i)} [(X(\zeta_2), Y(\zeta_2)) \in dy \times \{l\}] & \\ \times \mathbb{P}_{y,l} \left[ \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X(t), Y(t)) dt = \bar{f} \right] &= \\ \mathbb{P}_{(y,l)} \left[ \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X(t), Y(t)) dt = \bar{f} \right], & \end{aligned}$$

la última línea se obtuvo del uso de la distribución estacionaria. Con lo cual se muestra que al iniciar en un punto arbitrario  $(x, i)$  y una distribución inicial arbitraria es asintóticamente equivalente a iniciar con la distribución inicial siendo igual a la distribución estacionaria. Por lo que (157) es válido para todo  $(x, i) \in \mathbb{R}^r \times S$ . Con lo cual se concluye la prueba del teorema.

Q.E.D.

Como consecuencia del **Teorema 2.5.3** tenemos el siguiente resultado:

**Corolario 2.5.4:** *Supóngase que se satisfacen las condiciones del Teorema 2.5.3 y sea  $u(t, x, i)$  la solución del siguiente problema de Cauchy:*

$$\begin{cases} \frac{\partial u(t, x, i)}{\partial t} = \mathcal{L}u(t, x, i), & t > 0, (x, i) \in \mathbb{R}^r \times S, \\ u(0, x, i) = f(x, i), & (x, i) \in \mathbb{R}^r \times S, \end{cases} \quad (168)$$

donde  $f(\cdot, i) \in C_b(\mathbb{R}^r)$  para todo  $i \in S$ . Si  $T \rightarrow \infty$ , entonces

$$\frac{1}{T} \int_0^T u(t, x, i) dt \rightarrow \sum_{i=1}^N \int_{\mathbb{R}^r} f(x, i) \mu(x, i) dx. \quad (169)$$

**Demostración:**

En virtud del **Lema 2.3.9**,  $u(t, x, i) = \mathbb{E}_{(x,i)} [f(X(t), Y(t))]$ . Por lo que

$$\frac{1}{T} \int_0^T u(t, x, i) dt = \mathbb{E}_{(x,i)} \left[ \frac{1}{T} \int_0^T f(X(t), Y(t)) dt \right]. \quad (170)$$

Por otro lado, (157) implica que

$$\frac{1}{T} \int_0^T f(X(t), Y(t)) dt \xrightarrow{T \rightarrow \infty} \sum_{i=1}^N \int_{\mathbb{R}^r} f(x, i) \mu(x, i) dx, \text{ c.s.}$$

con respecto a la probabilidad  $\mathbb{P}_{(x,i)}$ . Por lo que la ecuación (169) se sigue del Teorema de convergencia dominada (véase **Teorema 1.4.15**).

Q.E.D.

Ahora, vamos a diseñar los controles adecuados para que la difusión cambiante resultante sea recurrente positiva y, por ende, ergódica. Considérese el proceso de difusión cambiante definido por (28) con transición del proceso de salto dada por (29) y condiciones iniciales  $(X(0), Y(0)) = (x, \alpha)$ . Considérese también el sistema de control asociado:

$$dX(t) = b(X(t), Y(t))dt + B(Y(t))u(X(t), Y(t))dt + \sigma(X(t), Y(t))dw(t), \quad (171)$$

donde  $B(i) \in \mathbb{R}^{r \times r}$ ,  $i \in S$  son matrices constantes y  $u(\cdot, \cdot) : \mathbb{R}^r \times S \rightarrow \mathbb{R}^r$  denota el control de retroalimentación.

Es bien sabido que el proceso de difusión cambiante (28) puede representarse a través de los siguientes  $N$  procesos de difusión:

$$dX_i(t) = b(X(t), i)dt + \sigma(X(t), i)dw(t), \quad i \in S, \quad (172)$$

agrupados por el componente discreto  $Y(t)$  de acuerdo a la transición especificada por (29). Casi siempre el sistema es observable sólo al operar en ciertas fases pero no en todas. Por ende, es usual descomponer el espacio de fases discreto  $S$  en dos subconjuntos disjuntos  $S_1$  y  $S_2$  tal que  $S = S_1 \cup S_2$ , de forma que para toda fase  $i \in S_2$  el proceso (172) no puede ser estabilizado por el control de retroalimentación, pero sí puede ser estabilizado para todo  $i \in S_1$ . Por ello consideraremos el control de retroalimentación de la siguiente forma:

$$u(X(t), Y(t)) = -L(Y(t))X(t),$$

donde para todo  $i \in S$ ,  $L(i) \in \mathbb{R}^{r \times r}$  es una matriz constante. Además, si  $i \in S_2$ ,  $L(i) = 0$ , por lo que (171) se reescribe como:

$$dX(t) = [b(X(t), Y(t)) - B(Y(t))L(Y(t))X(t)] dt + \sigma(X(t), Y(t))dw(t). \quad (173)$$

Asumiendo las condiciones **(A1.5.5)** y **(A1.5.6)**, es decir, suponemos que el sistema no lineal es localmente linealizable en una vecindad de  $\infty$ , entonces tenemos el siguiente teorema al aplicar el **Teorema 2.4.23** a (173).

**Teorema 2.5.5:** *Si para todo  $i \in S_1$ , existe una matriz constante  $L(i) \in \mathbb{R}^{r \times r}$  tal que:*

$$\begin{aligned}
& \sum_{i \in S_1} \pi_i \lambda_{\max} \left( b(i) - B(i)L(i) + b^*(i) - L^*(i)B^*(i) + \sum_{j=1}^d \sigma_j(i)\sigma_j^*(i) \right) \\
& + \sum_{i \in S_2} \pi_i \lambda_{\max} \left( b(i) + b^*(i) + \sum_{j=1}^d \sigma_j(i)\sigma_j^*(i) \right) < 0,
\end{aligned} \tag{174}$$

entonces el sistema de procesos de difusión cambiante controlado (171) es positivo recurrente.

El **Teorema 2.5.5** asegura que bajo condiciones simples, existe una variedad de elecciones para las matrices  $L(i)$ ,  $i \in S_1$  con el objeto de dotar al proceso de difusión cambiante (28) de la propiedad de recurrencia positiva. Resumiremos la discusión anterior en el siguiente teorema. Por ejemplo, considérese el conjunto de matrices

$$L(i) = \theta_i I, \quad i \in S_1, \tag{175}$$

de forma que  $I$  es la identidad de dimensión  $r \times r$  y  $\theta_i$  son constantes por determinar. Por lo que se sigue que para  $i \in S_1$ , se tiene que:

$$\begin{aligned}
& \lambda_{\max} \left( b(i) - B(i)L(i) + b^*(i) - L^*(i)B^*(i) + \sum_{j=1}^r \sigma_j(i)\sigma_j^*(i) \right) \\
& \leq \lambda_{\max} \left( b(i) + b^*(i) \sum_{j=1}^r \sigma_j(i)\sigma_j^*(i) \right) - \theta_i \lambda_{\min} (B(i) + B^*(i)).
\end{aligned} \tag{176}$$

Supóngase que para algún  $i \in S_1$  la matriz simétrica  $B(i) + B^*(i)$  es definida positiva. Por ende  $\lambda_{\min} (B(i) + B^*(i)) > 0$ . Por lo que si  $\theta_i > 0$  es lo suficientemente grande y  $\theta_j = 0$  para algún  $j \neq i$ , tendremos por (176) que el lado izquierdo de la ecuación (174) es menor que 0. Por lo que el sistema controlado (171) es recurrente positivo. Resumiremos lo anteriormente dicho en el siguiente teorema.

**Teorema 2.5.6:** *Si para alguna  $i \in S_1$  la matriz simétrica  $B(i) + B^*(i)$  es positiva definida, entonces existe un control de retroalimentación  $u(\cdot, \cdot)$  de forma que el proceso de difusión cambiante controlado (171) es positivo recurrente.*

Para finalizar esta sección se presentará una serie de simulaciones programadas en R para evidenciar algunos patrones de comportamiento de los procesos de difusión cambiante, en este caso el análisis de ergodicidad.

**Ejemplo 2.5.7:** *En este ejemplo ocuparemos el Teorema 2.5.5 y el Teorema 2.5.6 con el objeto de estabilizar (en el sentido débil) un proceso de difusión cambiante. Considérese el siguiente proceso bidimensional (en el componente continuo) que satisface que*

$$\begin{cases} dX(t) = b(X(t), Y(t))dt + \sigma(X(t), Y(t))dw(t), \\ \mathbb{P}[Y(t + \Delta) = j | Y(t) = i, X(s), Y(s), s \leq t] = q_{ij}(X(t))\Delta + o(\Delta), \end{cases} \tag{177}$$

al lado de su sistema de control

$$\begin{cases} dX(t) = b(X(t), Y(t))dt + B(Y(t))u(X(t), Y(t))dt + \sigma(X(t), Y(t))dw(t), \\ \mathbb{P}[Y(t + \Delta) = j | Y(t) = i, X(s), Y(s), s \leq t] = q_{ij}(X(t))\Delta + o(\Delta), \end{cases} \tag{178}$$

respectivamente. Supóngase que

$$\begin{aligned}
b(x_1, x_2, 1) &= \begin{bmatrix} x_1 + x_2 \\ 2x_2 \end{bmatrix}, \\
\sigma(x_1, x_2, 1) &= \begin{bmatrix} \sqrt{3 + 2x_2^2} & 5 \\ 0 & \sqrt{4 + x_2^2} - \sin(x_1 x_2) \end{bmatrix}, \\
b(x_1, x_2, 2) &= \begin{bmatrix} 2x_1 + x_2 \\ -x_1 + 3x_2 \end{bmatrix}, \\
\sigma(x_1, x_2, 2) &= \begin{bmatrix} -2 + \sqrt{1 + x_2^2} & 0 \\ -3 & \sqrt{2 + x_1^2} \end{bmatrix},
\end{aligned}$$

y

$$B(1) = \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}, B(2) = \begin{bmatrix} 2 & -3 \\ 1 & 3 \end{bmatrix}.$$

La matriz generadora de  $Y(\cdot)$  está dada por

$$Q(x) = \begin{bmatrix} -2 - \frac{\sin(x_1) \cos(x_1^2) + \sin(x_2^2)}{1 + \sqrt{x_1^2 + x_2^2}} & 2 + \frac{\sin(x_1) \cos(x_1^2) + \sin(x_2^2)}{1 + \sqrt{x_1^2 + x_2^2}} \\ 1 - \frac{2x_2 - \cos(x_1)}{3 + x_1^2 + x_2^2} & -1 + \frac{2x_2 - \cos(x_1)}{3 + x_1^2 + x_2^2} \end{bmatrix}.$$

Por lo que el proceso de difusión cambiante (177) tiene dos procesos asociados:

$$dX(t) = b(X(t), 1)dt + \sigma(X(t), 1)dw(t), \quad (179)$$

y

$$dX(t) = b(X(t), 2)dt + \sigma(X(t), 2)dw(t), \quad (180)$$

saltando del uno al otro de acuerdo al movimiento del componente discreto  $Y(t)$ , donde  $w(t) = [w_1(t), w_2(t)]^t$  es un movimiento Browniano bidimensional estándar. Supóngase que el sistema es observable cuando el componente discreto  $Y(\cdot)$  se encuentra en el estado 2. Los cálculos exactos (en virtud de la Sección 3.8 de [27]) permiten verificar que (179) y (180) son procesos transitivos. Véase las trayectorias de simuladas de los procesos.

## Ejemplo de Ergodicidad

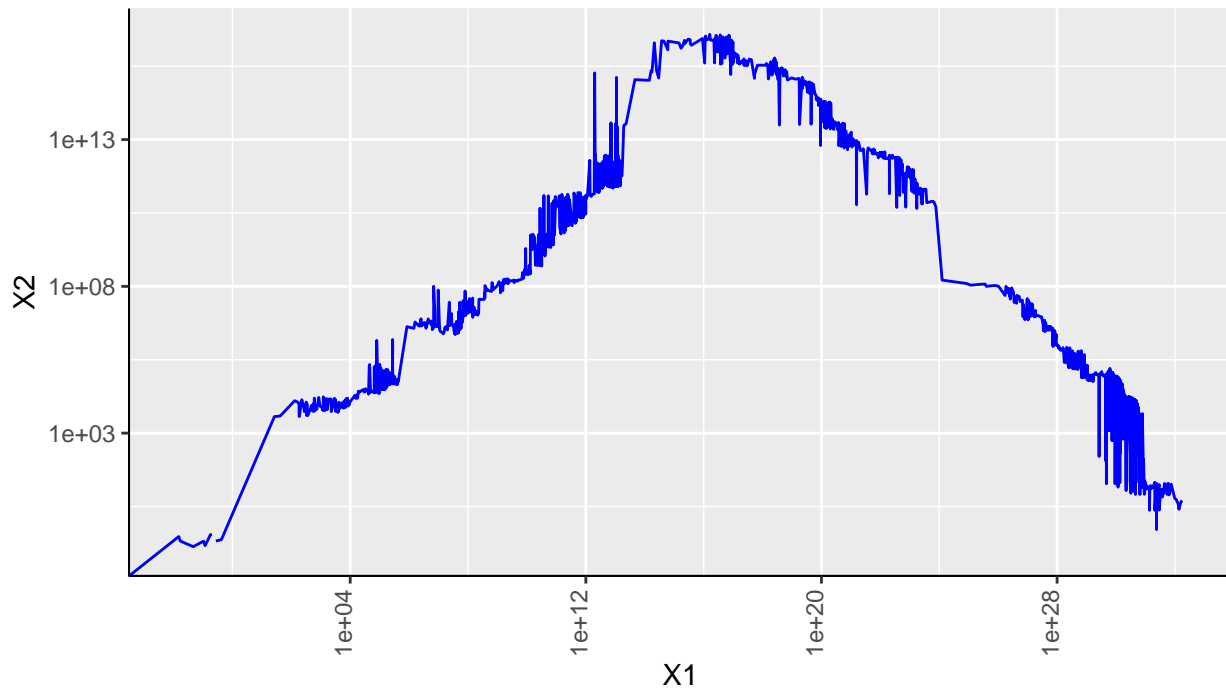


Figura 6: Trayectoria simulada del proceso (179) con condiciones iniciales  $(x_1, x_2) = (0, 0)$ , simulado por 40 segundos.

## Ejemplo de Ergodicidad

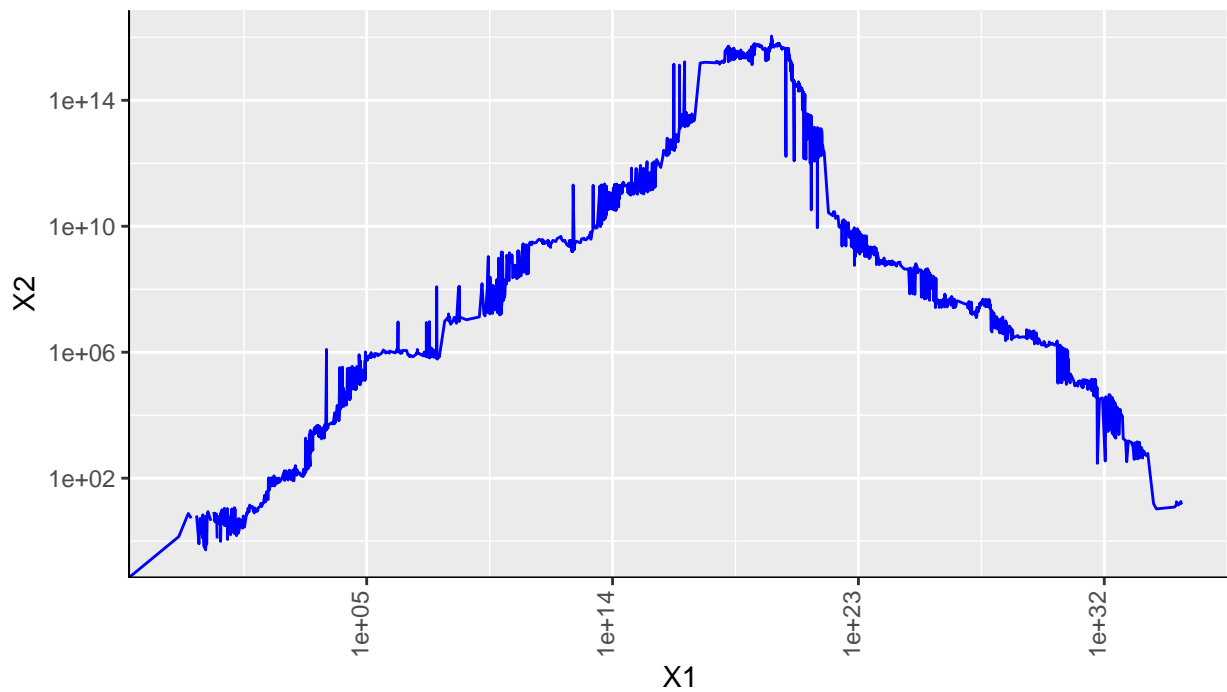


Figura 7: Trayectoria simulada del proceso (180) con condiciones iniciales  $(x_1, x_2) = (0, 0)$ , simulado por 40 segundos.

Ocuparemos el **Teorema 2.4.20** para verificar que el proceso (177) es transitivo. Para lograrlo, obtendremos las matrices  $b(i), \sigma_j(i)$  para  $i, j = 1, 2$  y también  $\hat{Q}$  en la condición **(A1.5.6)**:

$$\begin{aligned} b(1) &= \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \sigma_1(1) = \begin{bmatrix} 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}, \sigma_2(1) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \\ b(2) &= \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}, \sigma_1(2) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \sigma_2(2) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \end{aligned}$$

y

$$\hat{Q} = \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix}.$$

Por lo que la distribución estacionaria es  $\pi = [\frac{1}{3} \quad \frac{2}{3}]$ , y también tendremos que

$$\begin{aligned} \lambda_{\min}(b(1) + b^*(1) + \sigma_1(1)\sigma_1^*(1) + \sigma_2(1)\sigma_2^*(1)) - \frac{1}{2} \sum_{j=1}^2 [\rho(\sigma_j(1) + \sigma_j^*(1))]^2 &= 0.3820, \\ \lambda_{\min}(b(2) + b^*(2) + \sigma_1(2)\sigma_1^*(2) + \sigma_2(2)\sigma_2^*(2)) - \frac{1}{2} \sum_{j=1}^2 [\rho(\sigma_j(2) + \sigma_j^*(2))]^2 &= 4. \end{aligned}$$

Por lo que el **Teorema 2.4.20** implica que el proceso (177) es transitivo. Lo cual se puede verificar al observar la Figura 7.

Por suposición, la difusión cambiante es observable cuando el componente  $Y(\cdot)$  se encuentra en el estado 2. Nótese que

$$B(2) + B^*(2) = \begin{bmatrix} 4 & -2 \\ -2 & 6 \end{bmatrix},$$

es simétrica y positiva definida. Por ende y ocupando el **Teorema 2.5.6** se cumple que (178) es estabilizable en el sentido débil. De hecho, si consideramos  $\theta_1 = 0$  y  $\theta_2 = -4$  en (175), entonces se tiene que

$$\begin{aligned} \pi_2 \lambda_{\max}(b(2) + b^*(2) - 4B(2) - 4B^*(2) + \sigma_1(2)\sigma_1^*(2) + \sigma_2(2)\sigma_2^*(2)) \\ + \pi_1 \lambda_{\max}(b(1) + b^*(1) + \sigma_1(1)\sigma_1^*(1) + \sigma_2(1)\sigma_2^*(1)) &= -1.7647 < 0. \end{aligned}$$

Por lo tanto el sistema controlado (178) es recurrente positivo bajo el siguiente control de retroalimentación

$$u(x, 1) = 0 \text{ y } u(x, 2) = -4x. \quad (181)$$

### 3 Capítulo III: Análisis espectral de procesos de difusión cambiante

Este último Capítulo está enfocado en el análisis espectral de un proceso de difusión cambiante. Empezando por describir la notación que se ocupa en el Capítulo II sobre las principales propiedades del proceso, mientras que en el Capítulo II se ocupa una notación escalar, aquí ocuparemos una notación matricial para facilitar el subsecuente desarrollo del análisis espectral. Luego definimos y desarrollamos las *funciones propias o eigenfunciones* del operador infinitesimal del proceso (véase (36)) junto con las *ecuaciones de retroceso y evolución del proceso* (análogo a la Sección 1.2). Después se lleva a cabo un método para obtener la función de densidad de transición matricial del proceso en términos de las eigenfunciones del operador infinitesimal  $\mathcal{L}$ . Luego analizamos una serie de ecuaciones diferenciales que satisfacen los operadores asociados al proceso, todo para obtener un resultado que nos permita obtener la distribución estacionaria (análogo a (17)) de un proceso de difusión cambiante en términos de las eigenfunciones de  $\mathcal{L}$ . Para finalizar este capítulo se estudia un ejemplo de un proceso de difusión cambiante relacionado con el modelo de Wright-Fisher, se lleva a cabo el correspondiente análisis espectral de dicho y se incluyen varios gráficos basados en simulaciones de tal proceso.

#### 3.1 El operador infinitesimal

En este capítulo trabajaremos con un proceso  $(X(t), Y(t)) = (X_t, Y_t)$  que satisface (28) y (29) y cuyo componente continuo  $X(t)$  es unidimensional, es decir,  $\forall t \geq 0 X(t) \in \mathbb{R}$ . Se ocuparán como referencias principales a [12] y [36].

Volvemos a enunciar algunas definiciones y resultados que se mencionaron en el Capítulo II:

Considérese el operador de transición  $T_t f$  para  $f(x) \in \mathbb{B}(\mathbb{S}^N)$ , (similar a como se define en la **Definición 2.1.2** sólo que aquí consideramos la expresión vectorial y no únicamente la  $i$ -ésima componente), es decir:

$$T_t f(x) = \mathbb{E}[f(X_t) | X_0 = x] := \int_{\mathbb{S}} P(t; x, y) f(y) dy. \quad (182)$$

Como consecuencia de la ecuación de Chapman-Kolmogorov que satisface la función  $P(s + t; x, y)$ , resulta que la familia de operadores de transición  $\{T_t : t > 0\}$  satisface la propiedad de semigrupo  $T_{t+s} = T_s T_t$ , por ende también se satisface la propiedad de conmutatividad del operador  $T_t$ .

De manera similar, también ocuparemos el operador infinitesimal del proceso  $(X(t), Y(t))$ , definido en (36) para cualquier función vectorial suave  $f(x)$ , de forma que:

$$(\mathcal{L}f)(x) = \lim_{h \rightarrow 0} \frac{1}{h} [T_h f(x) - f(x)] = \lim_{h \rightarrow 0} \frac{1}{h} \left[ \int_{\mathbb{S}} P(h; x, y) f(y) dy - f(x) \right]. \quad (183)$$

Ahora, vamos a reescribir el operador  $\mathcal{L}$  bajo condiciones generalmente supuestas en la teoría de procesos de Markov a tiempo continuo:

- Para todo  $\epsilon > 0$  se satisface que:

$$\lim_{h \rightarrow 0} \frac{1}{h} P(h; x, y : |y - x| > \epsilon) = \mathbb{0}, \quad (184)$$

donde  $\mathbb{0}$  es la matriz nula.

- Existen dos funciones matriciales  $A(x), B(x) : x \in \mathbb{S}$  tales que para todo  $\epsilon > 0$ :

$$\lim_{h \rightarrow 0} \frac{1}{h} \left[ \int_{y: |y-x| \leq \epsilon} (y-x)^2 P(h; x, y) dy \right] = \frac{1}{2} A(x) = \frac{1}{2} \begin{bmatrix} \sigma^2(x, 1) & 0 & 0 & \dots & 0 \\ 0 & \sigma^2(x, 2) & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & \sigma^2(x, N) \end{bmatrix}, \quad (185)$$

$$\lim_{h \rightarrow 0} \frac{1}{h} \left[ \int_{y: |y-x| \leq \epsilon} (y-x) P(h; x, y) dy \right] = B(x) = \begin{bmatrix} b(x, 1) & 0 & 0 & \dots & 0 \\ 0 & b(x, 2) & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & b(x, N) \end{bmatrix}. \quad (186)$$

- Como se mencionó al principio del Capítulo II, existe una función matricial  $Q(x)$  de tal forma que

$$\lim_{h \rightarrow 0} \frac{1}{h} [P(h; x, S) - I_{N \times N}] = Q(x). \quad (187)$$

Véase (27).

Con el objeto de facilitar la notación y desarrollo subsecuente, a diferencia del Capítulo II, aquí ocuparemos una notación matricial-vectorial del operador infinitesimal.

Tal y como se demuestra en [4], bajo estas condiciones, el operador  $\mathcal{L}$  aplicado a  $f(x)$  adquiere la siguiente forma:

$$(\mathcal{L}f)(x) = \frac{1}{2} A(x) f''(x) + B(x) f'(x) + Q(x) f(x). \quad (188)$$

Y las matrices  $A(x), B(x)$  son matrices diagonales (véase [4]), de forma que denotaremos sus entradas diagonales por las funciones  $\sigma^2(x, i)$  y  $b(x, i), i = 1, \dots, N$ , respectivamente, así las ecuaciones (36) y (188) son equivalentes, (siendo que la primera ecuación es la expresión de cada entrada y la segunda es la expresión vectorial). De igual forma, también se sigue de la definición de  $Q(x) = [q_{ij}(x)]$  que para  $x \in \mathbb{S}$  la matriz satisface las condiciones de la matriz de intensidad de un proceso de Markov, es decir:

$$q_{ij}(x) = \begin{cases} \geq 0 : i \neq j \\ -\sum_{j \neq i} q_{ij}(x) \leq 0 : i = j \end{cases}. \quad (189)$$

En la siguiente sección, vamos a enlistar una serie de definiciones y resultados con respecto al operador  $\mathcal{L}$ , los cuales van a constituir la base para llevar a cabo un análisis espectral de los procesos de difusión cambiante.

### 3.2 Ecuaciones matriciales de retroceso y de evolución del proceso

Sea  $f(x) \in \mathbb{B}(\mathbb{S}^N)$  doblemente diferenciable, entonces la media definida a través del operador  $T_t f(x)$  satisface la *ecuación diferencial de retroceso* (análogo a (7)):

$$\frac{\partial}{\partial t} T_t f(x) = \frac{\partial}{\partial t} \mathbb{E}[f(X_t) | X_0 = x] = \mathcal{L} T_t f(x), \quad (190)$$

bajo condición inicial  $T_0 f(x) = f(x)$  (esta condición se verifica gracias a la suposición (187)). Lo cual es verificable al ocupar la definición del operador  $\mathcal{L}$  (183) aplicada a la función  $f(x) = T_t f(x)$  y la ecuación de Chapman-Kolmogorov (31).



Como consecuencia de lo anteriormente mencionado, la función de densidad de transición  $P(t; x, y)$  también satisface la *ecuación de retroceso de Kolmogorov* para todo  $t > 0, x, y \in \mathbb{S}$ :

$$\frac{\partial}{\partial t} P(t; x, y) = \frac{1}{2} A(x) \frac{\partial^2}{\partial x^2} P(t; x, y) + B(x) \frac{\partial}{\partial x} P(t; x, y) + Q(x) P(t; x, y). \quad (191)$$

Por otro lado, la *ecuación matricial de evolución o de Fokker-Planck* para  $P(t; x, y)$  puede considerarse la ecuación dual de (191). Para obtenerla, primero vamos a definir para toda función vectorial integrable  $g$  el siguiente operador:

$$T_t^* g(y) = \int_{\mathbb{S}} P^*(t; x, y) g(x) dx, \quad (192)$$

el cual corresponde al operador adjunto de  $T_t$  en el sentido de que:

$$\begin{aligned} \langle T_t f, g \rangle &= \int_{\mathbb{S}} g^*(x) \left( \int_{\mathbb{S}} P(t; x, y) f(y) dy \right) dx = \\ &= \int_{\mathbb{S}} \int_{\mathbb{S}} g^*(x) P(t; x, y) f(y) dx dy = \int_{\mathbb{S}} \left( \int_{\mathbb{S}} P^*(t; x, y) g(x) dx \right)^* f(y) dy = \langle f, T_t^* g \rangle. \end{aligned} \quad (193)$$

Nótese que  $\langle f, g \rangle = \int_{\mathbb{S}} g^*(x) f(x) dx$  es un producto interior real para cualesquiera dos funciones vectoriales.

En el caso de que  $f$  fuese continua, doblemente diferenciable y de soporte compacto contenido en  $\mathbb{S}$ , al derivar con respecto a  $t$  en (193) e intercambiar el orden de derivación e integración se obtiene (valiéndose de (190)), el siguiente resultado:

$$\left\langle f, \frac{\partial}{\partial t} T_t^* g \right\rangle = \left\langle \frac{\partial}{\partial t} T_t f, g \right\rangle = \langle T_t \mathcal{L} f, g \rangle = \langle \mathcal{L} f, T_t^* g \rangle = \langle f, \mathcal{L}^* T_t^* g \rangle. \quad (194)$$

Con  $\mathcal{L}^*$  nos referimos al operador adjunto de  $\mathcal{L}$  con respecto al producto interior  $\langle \cdot, \cdot \rangle$ , el cual está dado por:

$$(\mathcal{L}^* g)(y) = \frac{1}{2} \frac{\partial^2}{\partial y^2} (A^*(y) g(y)) - \frac{\partial}{\partial y} (B^*(y) g(y)) + Q^*(y) g(y). \quad (195)$$

Ocupando (194), se obtiene la siguiente igualdad, la cual corresponde a la *ecuación diferencial de evolución* (análogo a (10))

$$\frac{\partial}{\partial t} T_t^* g(y) = \mathcal{L}^* T_t^* g(y). \quad (196)$$

De forma tal que la función de densidad de transición  $P(t; x, y)$  satisface que

$$\frac{\partial}{\partial t} P^*(t; x, y) = \frac{1}{2} \frac{\partial^2}{\partial y^2} (A^*(y) P^*(t; x, y)) - \frac{\partial}{\partial y} (B^*(y) P^*(t; x, y)) + Q^*(y) P^*(t; x, y). \quad (197)$$

Y ocupando el adjunto:

$$\frac{\partial}{\partial t} P(t; x, y) = \frac{1}{2} \frac{\partial^2}{\partial y^2} (P(t; x, y) A(y)) - \frac{\partial}{\partial y} (P(t; x, y) B(y)) + P(t; x, y) Q(y). \quad (198)$$

Importante notar que la *ecuación diferencial de evolución* (196) tiene los coeficientes del operador infinitesimal multiplicados por la derecha, mientras que *ecuación diferencial de retroceso* tiene los coeficientes del operador infinitesimal multiplicados por la izquierda (191).

### 3.3 Representación espectral de la función de densidad de transición matricial

En esta sección, se mostrará un método para obtener una fórmula para calcular  $P(t; x, y)$  en términos de las funciones propias matriciales del operador infinitesimal  $\mathcal{L}$ . Nuevamente, volvemos a mencionar algunas definiciones y resultados del Capítulo I.

Sea  $L_W^2(\mathbb{S}, \mathbb{C}^{N \times N})$  el conjunto de todas las funciones matriciales  $F(x)$  las cuales satisfacen:

$$\langle F, F \rangle_W = \int_{\mathbb{S}} F^*(x) dW(x) F(x) dx < \infty,$$

tal que  $\langle \cdot, \cdot \rangle_W$  es el producto (24) y  $dW$  es una matriz peso con densidad  $W$  (véase **Definición 1.3.6**) absolutamente continua con respecto a la medida de Lebesgue (véase [13] y (23)).

Nótese que si  $F(x) \in L_W^2(\mathbb{S}, \mathbb{C}^{N \times N})$ , al decir  $\langle F, F \rangle_W < \infty$  nos referimos a que entrada a entrada la integral matricial es finita. Es más, este producto matricial  $\langle \cdot, \cdot \rangle_W$  no es un producto escalar en el sentido estricto, aun así posee propiedades similares a las de un producto interior escalar regular.

Por otro lado, es posible definir un producto escalar de funciones matriciales (véase (25)):

$$\langle\langle F, G \rangle\rangle_W := \text{Tr}(\langle F, G \rangle_W).$$

Por lo tanto, al darle al espacio  $L_W^2(\mathbb{S}, \mathbb{C}^{N \times N})$  la norma  $\|F\|_W = \sqrt{\text{Tr}(\langle F, F \rangle_W)}$ , lo estamos ajustando para que sea un espacio de Hilbert con producto interior  $\langle\langle F, G \rangle\rangle_W$  (véase **Definición 1.3.9**).

Supongamos la existencia de un conjunto numerable de funciones matriciales propias  $\{\Phi_n(x)\}_n$  del operador  $\mathcal{L}$  con valores propios hermitianos matriciales  $\{\Gamma_n\}_n \in \mathbb{C}^{N \times N}$ , de forma que  $\mathcal{L}\Phi_n(x) = \Gamma_n\Phi_n(x)$ . Supóngase también que  $\langle\Phi_n, \Phi_m\rangle_W = \delta_{nm} Id_{N \times N}$  para alguna matriz peso  $W$ .

En el caso de que el conjunto de todas las combinaciones lineales finitas de  $\{\Phi_n(x)\}_n$  sea completo en  $L_W^2(\mathbb{S}, \mathbb{C}^{N \times N})$  entonces cualquier  $F \in L_W^2(\mathbb{S}, \mathbb{C}^{N \times N})$  posee una expansión en series de Fourier de la forma:

$$F(x) = \sum_{n=0}^{\infty} \Phi_n(x) \langle F, \Phi_n \rangle_W. \quad (199)$$

Considérese también la siguiente superposición de  $F$  de la forma:

$$U_F(t, x) = \sum_{n=0}^{\infty} \Phi_n(x) e^{t\Gamma_n} \langle F, \Phi_n \rangle_W. \quad (200)$$

Nótese que  $e^{t\Gamma_n}$  está definido a partir de la expansión en series de Taylor, es decir  $e^{t\Gamma_n} = \sum_{n=0}^{\infty} \frac{t^n (\Gamma_n)^n}{n!}$ .

De igual forma, obsérvese que la función  $U_F(t, x)$  satisface la ecuación diferencial de retroceso:

$$\frac{\partial}{\partial t} U_F(t, x) = \mathcal{L}U_F(t, x),$$

con condición inicial  $U_F(0, x) = F(x)$ . Por ende, de existir unicidad para una clase suficientemente grande de funciones iniciales  $F$ , tendremos que:

$$U_F(t, x) = T_t F(x).$$

Lo anterior implica que

$$\int_S P(t; x, y) F(y) dy = \int_S \left( \sum_{n=0}^{\infty} \Phi_n(x) e^{t\Gamma_n} \Phi_n^*(x) W(y) \right) F(y) dy.$$

Por lo tanto,  $P(t; x, y)$  tiene la forma:

$$P(t; x, y) = \sum_{n=0}^{\infty} \Phi_n(x) e^{t\Gamma_n} \Phi_n^*(x) W(y). \quad (201)$$

Cabe resaltar que al condicionar a que los valores propios de  $\mathcal{L}$  sean hermitianos estamos forzando al operador a ser auto-adjunto con respecto al producto interior, es decir  $F, G \in L_W^2(\mathbb{S}, \mathbb{C}^{N \times N})$ ,  $\langle \mathcal{L}F, G \rangle_W = \langle F, \mathcal{L}G \rangle_W$ .

Lo anterior implica que al ocupar la definición formal del operador adjunto a  $\mathcal{L}$  en (195), los coeficientes del operador infinitesimal  $\mathcal{L}$  están condicionados a satisfacer ciertas *ecuaciones de simetría* con respecto a la matriz peso  $W$  (con sus respectivas condiciones de frontera):

$$\begin{cases} A^*(x)W(x) = W(x)A(x), \\ B^*(x)W(x) = (W(x)A(x))' - W(x)B(x), \\ Q^*(x)W(x) = \frac{1}{2}(W(x)A(x))'' - (W(x)B(x))' + W(x)Q(x). \end{cases} \quad (202)$$

Estas ecuaciones de simetría fueron obtenidas en [13] al integrar por partes la ecuación  $\langle \mathcal{L}F, G \rangle_W = \langle F, \mathcal{L}G \rangle_W$ .

### 3.4 Ecuaciones diferenciales asociadas a algunos operadores

En esta sección analizaremos algunas ecuaciones diferenciales asociadas a operadores relacionados con las distribuciones invariantes y la recurrencia de los procesos de difusión cambiante.

El tiempo de llegada de un proceso unidimensional  $\{X_t, 0 \leq t < \zeta\}$  a un estado  $z \in \mathbb{S} = (a, b)$  se define como  $T_z = \inf\{t \geq 0 : X_t = z\}$  o  $T_z = +\infty$  en el caso donde  $X_t \neq z : 0 \leq t < \zeta$ .

De forma similar, definimos  $T^* = \min\{T_c, T_d\}$  como el tiempo de llegada a  $c$  o  $d$ , es decir el primer momento  $t$  donde  $X_t = c$  o  $X_t = d$ , para algunos  $c, d \in \mathbb{S}$ . Para procesos que inicien en  $X_0 \in (c, d)$ ,  $T^*$  equivale al tiempo de salida del intervalo  $(c, d)$  (véase **Definición 1.2.8**).

Dado un proceso de difusión cambiante  $(X(t), Y(t))$ , decimos que este es regular en el interior de  $\mathbb{S}$  si

$$\mathbb{P} [T_y < +\infty, Y_{T_y} = j | X_0 = x, Y_0 = i] > 0,$$

para cualesquiera  $x, y \in \text{int}(\mathbb{S})$  e  $i, j$  fases (véase **Definición 2.2.2**).

**Definición 3.4.1** Definimos la función matricial  $U(x) = [U_{ij}(x)]$ , de forma tal que:

$$U_{ij}(x) = \mathbb{P} [T_d < T_c, Y_{T_d} = j | X_0 = x, Y_0 = i], x \in (c, d), \quad (203)$$

en otras palabras, iniciando en  $x \in (c, d)$  y en la fase  $i$ ,  $U_{ij}(x)$  es la probabilidad de que proceso llegue a  $d$  antes de llegar a  $c$  y que al mismo tiempo el proceso esté en la fase  $j$ .

**Teorema 3.4.2:** La matriz  $U(x)$  previamente definida en (203), es doblemente diferenciable para toda  $x \in (c, d)$  y satisface la siguiente ecuación diferencial:

$$\frac{1}{2}A(x)U''(x) + B(x)U'(x) + Q(x)U(x) = 0, \quad x \in (c, d), \quad U(c) = 0, \quad U(d) = I. \quad (204)$$

**Demostración:**

La prueba es básicamente la misma para el caso unidimensional, (véase [26] p. 193, Proposición 9.1 y Capítulo V de [6]).

Q.E.D

**Definición 3.4.3:** Definimos la función matricial  $R_{x,y} = [R_{x,y}]_{ij}$ , de forma tal que:

$$[R_{x,y}]_{ij} = \mathbb{P}[X_t = y, Y_t = j | X_0 = x, Y_0 = i], \quad x, y \in \mathbb{S}, \quad t \geq 0,$$

en otras palabras, la probabilidad de que al empezar en el estado  $x$  y en la fase  $i$ , el proceso llegue en algún momento al estado  $y$ , y en la fase  $j$ .

Nótese que si  $y > x$ , entonces  $R_{x,y}$  será la solución de la ecuación diferencial (204) sustituyendo  $c \rightarrow a$ , y  $d \rightarrow y$ , mientras que si  $x > y$ , entonces  $R_{x,y}$  será la solución de la ecuación diferencial (204) sustituyendo  $c \rightarrow b$ , y  $d \rightarrow y$ .

**Definición 3.4.4** Un estado  $y$  es recurrente si  $R_{x,y}[e_N] = e_N$  para todo  $x \in \mathbb{S}$ , donde  $e_N$  es un vector columna de dimensión  $N$  cuyas entradas son iguales a 1. Decimos que  $y$  es transitivo en caso contrario. Si todos los estados en  $S$  son recurrentes, diremos entonces que el proceso de Markov bivariado es recurrente.

Nótese que la **Definición 2.4.1** nos dice que para el caso particular de que  $U = X \times \{j\}$ , donde  $X \in \mathbb{S}$  es arbitrario y  $j \in S$  es fijo, y  $\sigma_U^{x,i} = \inf\{t \geq 0 : (X^{x,i}(t), Y^{x,i}(t)) \in U\}$ , el proceso  $(X^{x,i}(t), Y^{x,i}(t))$  es recurrente con respecto a  $U$  si y solo si  $\mathbb{P}[\sigma_U^{x,i} < \infty] = 1$ . Por lo que

$$\mathbb{P}[\sigma_U^{x,i} < \infty] = 1 \iff \mathbb{P}[\sigma_{X \times \{j\}}^{x,i} < \infty] = 1 \iff \exists t \in \mathbb{R}_+ : \mathbb{P}[X_t = X, Y_t = j | X_0 = x, Y_0 = i] = 1,$$

es decir, la **Definición 3.4.4** y la **Definición 2.4.1** son equivalentes.

**Definición 3.4.5:** Sea  $G(x)$  una función matricial continua, definimos  $V(x)$  como:

$$V(x) = \mathbb{E} \left[ \int_0^{T^*} G(X_s) ds | X_0 = x \right], \quad x \in (c, d), \quad (205)$$

donde  $\mathbb{E}$  es de la forma del operador de transición definido en (182).

**Teorema 3.4.6** La matriz  $V(x)$  previamente definida en (205), es doblemente diferenciable para toda  $x \in (c, d)$  y satisface la siguiente ecuación diferencial:

$$\frac{1}{2}A(x)V''(x) + B(x)V'(x) + Q(x)V(x) + G(x) = 0, \quad x \in (c, d), \quad V(c) = 0, \quad V(d) = 0. \quad (206)$$

**Demostración:**

Nuevamente, la prueba es básicamente la misma para el caso unidimensional, (véase [26] p. 193, Proposición 9.1 y Capítulo V de [6]).

Q.E.D

Nótese que en el caso de que

$$G(x) = e_N e_N^T = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix}, \quad (207)$$

entonces tendremos que la función  $V(x) = [V_{ij}(x)]$  será de la forma:

$$V_{ij}(x) = \mathbb{E}[T^*, Y_{T^*} = j | X_0 = x, Y_0 = i], \quad (208)$$

en otras palabras, empezando en el estado  $X_0 = x$  y en la fase  $Y_0 = i$ ,  $V_{ij}(x)$  el tiempo medio para alcanzar el estado  $c$  o el estado  $d$  y que al mismo tiempo el proceso se encuentre en la fase  $j$ .

**Definición 3.4.7** *Supóngase que el proceso  $(X(t), Y(t))$  es recurrente. Si sustituimos  $c \rightarrow a$  y  $d \rightarrow b$  en (206) del Teorema 3.4.6, decimos que  $(X(t), Y(t))$  es recurrente positivo si todas las entradas de  $V(x)$  en (208), con  $G(x) = e_N e_N^T$  (tal y como está definida en (207)) son finitas. De otro modo, diremos que  $(X(t), Y(t))$  es recurrente nulo.*

Nótese que en este caso, la **Definición 3.4.7** deriva directamente del **Teorema 2.4.21**, de forma que se sigue el Teorema para  $u(x, i) = V_{ij}(x) = \mathbb{E}[T^*, Y_{T^*} = j | X_0 = x, Y_0 = i]$ .

**Definición 3.4.8:** *Sea  $\psi(y) := (\psi_1(y), \psi_2(y), \dots, \psi_N(y))$  una función vectorial  $\psi : \mathbb{S} \rightarrow \mathbb{R}^N$ . Diremos que  $\psi$  es distribución estacionaria de  $(X(t), Y(t))$  si satisface que para todo  $t > 0$ :*

$$\psi(y) = \int_{\mathbb{S}} \psi(x) P(t; x, y) dx.$$

Nótese que, ocupando una metodología similar a la que ocupamos para la obtención de la *ecuación diferencial de evolución* en (192), tendremos que  $\psi(y)$  satisface que:

$$\frac{1}{2}(\psi(y)A(y))'' - (\psi(y)B(y))' + \psi(y)Q(y) = 0. \quad (209)$$

Además, con el objeto de tener una distribución de probabilidad, debemos imponer las siguientes condiciones sobre  $\psi(y)$ :

$$\begin{aligned} 0 \leq \psi_j(y) \leq 1, \quad j = 1, \dots, N \\ \left( \int_{\mathbb{S}} \psi(y) dy \right) e_N = 1. \end{aligned} \quad (210)$$

Si el operador infinitesimal  $\mathcal{L}$  es auto-adjunto con respecto a alguna matriz peso  $W$ , entonces existe una solución explícita a la ecuación (209) previamente definida, la cual es de la forma:

$$\psi(y) = c e_N^T W(y), \quad c > 0.$$

Nótese que ocupando la *tercera ecuación de simetría* en (202) tendremos la siguiente igualdad:

$$c e_N^T \left[ \frac{1}{2}(W(y)A(y))'' - (W(y)B(y))' + W(y)Q(y) \right] = c e_N^T Q^*(y)W(y) = 0,$$

lo último es consecuencia del hecho de que  $Q(y)e_N = 0$  para todo  $y$ . Finalmente, ocupando (210) tendremos el valor exacto de la constante  $c$ :

$$c = \left( \int_{\mathbb{S}} e_N^T W(y) e_N dy \right)^{-1}.$$

Por lo tanto, la fórmula explícita para obtener la distribución estacionaria de  $(X(t), Y(t))$ , (en el caso donde  $\mathcal{L}$  sea auto-adjunto con respecto  $W$ ), es:

$$\psi(y) = \left( \int_{\mathcal{S}} e_N^T W(y) e_N dy \right)^{-1} e_N^T W(y).$$

Resumiremos todo lo anteriormente mencionado en el siguiente teorema.

**Teorema 3.4.9:** *Si el operador infinitesimal  $\mathcal{L}$  es auto-adjunto con respecto a alguna matriz de peso  $W$  y el proceso  $(X(\cdot), Y(\cdot))$  es recurrente positivo, entonces existe una solución explícita a la ecuación (209) previamente definida, la cual es de la forma:*

$$\psi(y) = \left( \int_{\mathcal{S}} e_N^T W(y) e_N dy \right)^{-1} e_N^T W(y).$$

### 3.5 Ejemplo de análisis espectral de procesos de difusión cambiante

Para finalizar este capítulo se estudiará un ejemplo de análisis espectral de un proceso de difusión cambiante. En particular, nos enfocaremos en una variante del modelo de difusión de Wright-Fisher, estudiado en [12].

El modelo de difusión de Wright-Fisher que involucra sólo efectos de mutación toma en consideración una población de tamaño constante compuesta por dos tipos: A y B. Las tasas de intensidad de mutar de A a B y viceversa están dadas por las constantes positivas  $\frac{1+\beta}{2}$  y  $\frac{1+\alpha}{2}$  respectivamente (donde  $\beta, \alpha > -1$ ). Cuando el tamaño de la población tiende a infinito es bien sabido que este modelo puede describirse mediante un proceso de difusión cuyo espacio de estados es el intervalo unitario  $\mathcal{S} = [0, 1]$  (véase [26], página 177) con coeficientes de deriva y de difusión dados por

$$b(x) = \alpha + 1 - x(\alpha + \beta + 2), \quad \sigma^2(x) = 2x(1 - x), \quad \beta, \alpha > -1. \quad (211)$$

La variable  $x$  es la tasa del número de miembros de la población pertenecientes a la clase A. Estos coeficientes corresponden con los del modelo de difusión de Jacobi, por lo que la densidad de probabilidad puede describirse en términos de los polinomios de Jacobi sobre el intervalo  $[0, 1]$  (véase **Ejemplo 1.3.5**).

En esta sección estudiaremos una variación del modelo de difusión de Wright-Fisher donde los efectos de mutación están determinados por un proceso de Markov bivariado cuyo espacio de estados es el conjunto  $[0, 1] \times \{1, 2, \dots, N\}$ , de forma que se pueda llevar a cabo el análisis que se presentó en las secciones anteriores. En este caso, habrá  $N$  fases posibles con distintos coeficientes de deriva (el coeficiente de difusión siempre será el mismo). Además, aparecerá un parámetro  $k$  el cual mide la forma en la que el proceso se mueve a través de las distintas fases.

El ejemplo que analizaremos proviene de la teoría de representación de grupos, tiene que ver con funciones esféricas matriciales asociadas al espacio proyectivo complejo. Para más información consultar [22] y las referencias contenidas en este artículo.

En lo que sigue ocuparemos  $E_{ij}$  para denotar la matriz cuya entrada  $(i, j)$  es 1 y el resto es 0. Sea  $N \in \{1, 2, \dots\}$ ,  $\alpha, \beta > -1$  y  $0 < k < \beta + 1$ . Ahora, sea  $M$  la matriz triangular superior nilpotente dada por

$$M = \sum_{i=1}^{N-1} E_{i, i+1},$$

y sean  $J, \hat{J}$  las matrices diagonales determinadas por

$$J = \sum_{i=1}^N (N - i) E_{i, i}, \quad \hat{J} = \sum_{i=1}^N (i - 1) E_{i, i} = (N - 1)I - J.$$

Ahora, sea  $H$  la matriz diagonal dada por

$$H = \sum_{i=1}^N \omega_i E_{i,i},$$

donde

$$\omega_i = \binom{\beta - k + i - 1}{i - 1} \binom{N + k - i - 1}{N - i}.$$

Nótese que los pesos  $\omega_i$  corresponden a los pesos de los polinomios de Hahn con soporte en  $\{1, 2, \dots, N\}$  con parámetros  $\beta - k > -1$  y  $k - 1 > -1$  (véase [2]).

El par  $\{W, \mathcal{L}\}$ , tal que

$$W(x) = x^\alpha (1 - x)^\beta H x^J, x \in (0, 1), \quad (212)$$

y

$$\mathcal{L} = \frac{1}{2} A(x) \frac{d^2}{dx^2} + B(x) \frac{d}{dx} + Q(x) \frac{d^0}{dx^0}, \quad (213)$$

donde

$$\begin{aligned} A(x) &= 2x(1 - x)I, \quad B(x) = (\alpha - 1)I + J - x((\alpha + \beta + 2)I + J), \\ Q(x) &= \frac{1}{1 - x} \left( J((\beta - k + 1)I + \hat{J}) (M - 1) + x(\hat{J}(kI + J)(M^* - I)) \right), \end{aligned} \quad (214)$$

es un *par simétrico*, lo que implica que  $\mathcal{L}$  es auto-adjunto con respecto al peso  $W$  (todo con respecto al producto interior  $\langle \cdot, \cdot \rangle_W$  (véase inicio de la Sección 3.3)).

La matriz tridiagonal  $Q(x)$  puede escribirse como

$$Q(x) = \sum_{i=2}^N \mu_i(x) E_{i,i-1} - \sum_{i=1}^N (\lambda_i(x) + \mu_i(x)) E_{ii} + \sum_{i=1}^{N-1} \lambda_i(x) E_{i,i+1}, \quad (215)$$

donde

$$\begin{aligned} \lambda_i(x) &= \frac{1}{1 - x} (N - i)(i + \beta - k), \\ \mu_i(x) &= \frac{x}{1 - x} (i - 1)(N - i + k). \end{aligned} \quad (216)$$

Cabe resaltar que  $\lambda_i(x)$  y  $\mu_i(x)$  son variaciones  $x$ -dependientes de los coeficientes de la ecuación de diferencias de segundo orden que satisfacen los polinomios de Hahn (véase [2], página 346).

Nótese que  $A(x)$  y  $B(x)$  son matrices diagonales y las entradas diagonales de  $A(x)$  son positivas para todo  $x \in (0, 1)$ . Además de que  $\lambda_i(x)$  y  $\mu_i(x)$  son siempre positivas y que  $Q(x)e_N = 0$  para todo  $x \in (0, 1)$ . Por lo que  $Q(x)$  es la matriz de intensidad de una cadena de Markov a tiempo continuo con espacio de estados finito. Con lo cual se satisface la definición de un proceso de Markov bivariado.

Las eigenfunciones  $(\Phi_n(x))_n$  del operador  $\mathcal{L}$  (213) (conocidas también como las funciones esféricas matriciales asociadas al espacio proyectivo complejo) están relacionadas con los polinomios matriciales ortogonales definidos en [22]. Eso se debe a que posee la mayoría de las fórmulas estructurales necesarias en el subsecuente

análisis. Primero describiremos la relación entre el par simétrico  $\{W, \mathcal{L}\}$  y de su correspondiente contraparte  $\{\widetilde{W}, \widetilde{\mathcal{L}}\}$  dadas en [22]. Denótese por  $\Psi(x)$  a la siguiente matriz

$$\Psi(x) = e^{\hat{J}M^T} (1-x)^{\hat{J}} T,$$

donde  $T$  es la matriz triangular superior constante dada al principio de la Sección 7 de [22], es decir:

$$T = \sum_{i \leq j}^N (-1)^i \frac{(-j)_i}{(1-N)_i} \frac{(\alpha + \beta - k + j + 1)_i}{(\beta - k + 1)_i} E_{ij},$$

donde  $(a)_i$  denota el símbolo de Pochhammer, el cual se define como  $(a)_n = a(a+1) \cdots (a+n-1)$  para toda  $n > 0$ ,  $(a)_0 = 1$ .

Los pares  $\{W, \mathcal{L}\}$  y  $\{\widetilde{W}, \widetilde{\mathcal{L}}\}$  se relacionan de la siguiente manera:

$$\widetilde{W}(x) = \Psi^T(x)W(x)\Psi(x), \quad \widetilde{\mathcal{L}}F(x) = \Psi^{-1}(x)\mathcal{L}(\Psi(x)F(x)). \quad (217)$$

Ahora, ocupando la familia de polinomios ortogonales matriciales que se introdujo en [22], a los que denotaremos por  $(R_n(x))_n$ , entonces la familia  $(R_n(x))_n$  es un conjunto de eigenfunciones del operador diferencial  $\widetilde{\mathcal{L}}$ , es decir  $\widetilde{\mathcal{L}}R_n(x) = R_n(x)\Gamma_n$ , donde  $\Gamma_n$  es la matriz diagonal

$$\Gamma_n = -n^2 I - n((\alpha + \beta + N)I + \hat{J}) - \hat{J}((\alpha + \beta - k + 1)I + \hat{J}). \quad (218)$$

Esta familia posee propiedades especiales. En particular es fácil normalizarla dado que las normas  $\|R_n\|_{\widetilde{W}}^2$  dadas en la fórmula (7.3) de [22] son matrices diagonales. Las normas  $\|R_n\|_{\widetilde{W}}^2$  se expresan de la siguiente manera:

$$\begin{aligned} \|R_n\|_{\widetilde{W}}^2 &= \langle R_n, R_n \rangle_{\widetilde{W}} = \sum_{i=0}^{N-1} (-1)^i \frac{\Gamma(n+1)\Gamma(\beta+N)\Gamma(\alpha+N+i+1)(1-k-N-n)_i}{\binom{N-1}{i} \Gamma(N)\Gamma(\alpha+\beta+N+i+2n+1)} \\ &\quad \times \frac{(k+N-i-1)_n (\alpha+\beta+N+i+N)_n (\alpha+\beta-k+i+n+1)_N}{(\alpha+\beta-k+2i+n+1)(k)_n (\beta-k+1)_i (\beta+N)_n} E_{ii}, \end{aligned}$$

donde  $\Gamma(t)$  es la función Gamma estándar.

Por ende, defínase la familia

$$\Phi_n(x) = \Psi(x)R_n(x)\|R_n\|_{\widetilde{W}}^{-1}.$$

Por lo tanto, tendremos que esta familia es el conjunto de eigenfunciones del operador infinitesimal  $\mathcal{L}$ , es decir

$$\mathcal{L}\Phi_n(x) = \Phi_n(x)\Gamma_n,$$

ortonormales con respecto a  $W(x)$  (212) y  $\Gamma_n$  definidos en (218). Nótese que  $(\Phi_n(x))_n$  son además polinomios matriciales, pero no de grado  $n$  precisamente (para ser exactos, son de grado  $n+N-1$ ).

Por lo que la fórmula (201) nos dice que

$$P(t; x, y) = \sum_{n=0}^{\infty} \Phi_n(x) e^{t\Gamma_n} \Phi_n^T(y) W(y), \quad (219)$$



o (en términos del par  $\{\widetilde{W}, \widetilde{\mathcal{L}}\}$ )

$$P(t; x, y) = \Psi(x) \left( \sum_{n=0}^{\infty} R_n(x) e^{t\Gamma_n} \|R_n\|_{\widetilde{W}}^{-2} R_n^T(y) \widetilde{W}(y) \right) \Psi^{-1}(y).$$

La familia del proceso de difusión cambiante  $(X(t), Y(t))$  con espacio de estados  $[0, 1] \times \{1, 2, \dots, N\}$  puede describirse en términos de la ecuación diferencial estocástica

$$dX(t) = b(X(t), Y(t))dt + \sigma(X(t))dw(t), \quad (220)$$

donde

$$b(X(t), i) = \alpha + 1 + N - i - x(\alpha + \beta + 2 + N - i), \quad \sigma^2(x) = 2x(1 - x), \quad Y(t) = 1, \dots, N. \quad (221)$$

La manera en la que el proceso cambia de fases está determinada por el generador  $Q(x)$  (215). Nótese que  $Q(x)$  es tridiagonal, por lo que la transición entre fases sólo está permitida entre fases adyacentes (como un proceso de nacimiento y muerte).

El comportamiento de  $X(t)$  depende de la fase  $Y(t)$  y la transición entre fases y tiempos de espera en cada fase  $Y(t)$  dependen de la posición del componente continuo  $X(t)$ . Con respecto a la parte continua  $X(t)$ , nos interesa estudiar el comportamiento en los valores de frontera, en este caso 0 y 1. Por el contrario, para la parte discreta  $Y(t)$  nos interesa cómo el proceso se mueve a través de distintas fases, incluyendo los tiempos de espera entre fases.

El coeficiente de difusión  $\sigma^2(x)$  siempre es el mismo, pero los coeficientes de deriva  $b(X(t), i)$  dependen de la fase  $i$ . Para una fase  $i$ , el moverse a la siguiente fase  $i + 1$  significa disminuir en una unidad la intensidad de la mutación  $B \rightarrow A$  (en términos del parámetro  $\alpha$ ) y moverse a la fase anterior  $i - 1$  significa aumentar en una unidad la intensidad de la mutación  $B \rightarrow A$  (en términos del parámetro  $\alpha$ ). El caso límite es en la fase  $N$ , donde se recupera la difusión de Jacobi regular (211). La intensidad de la mutación  $A \rightarrow B$  (en términos del parámetro  $\beta$ ) no es afectada en estos coeficientes.

$\alpha=1, \beta=1, k=1.5, 50$  cambios

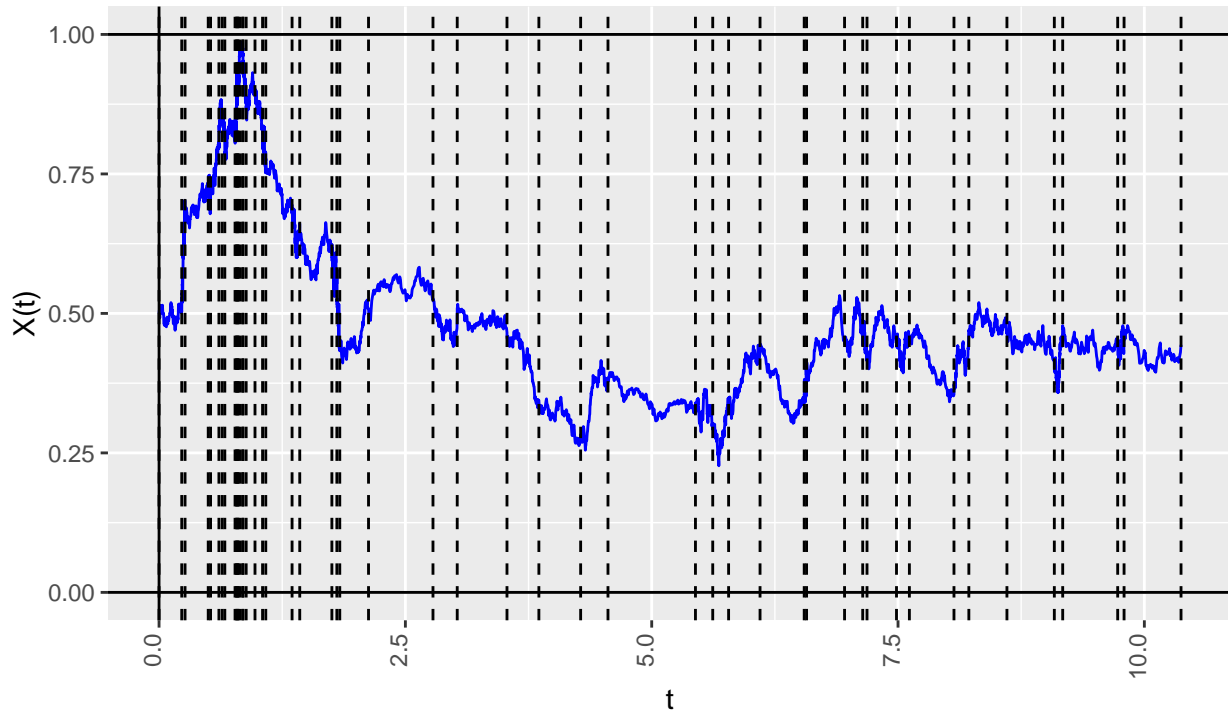


Figura 8.1 Simulación del proceso (220)-(221)

$\alpha=1, \beta=-0.4, k=0.3, 46$  cambios

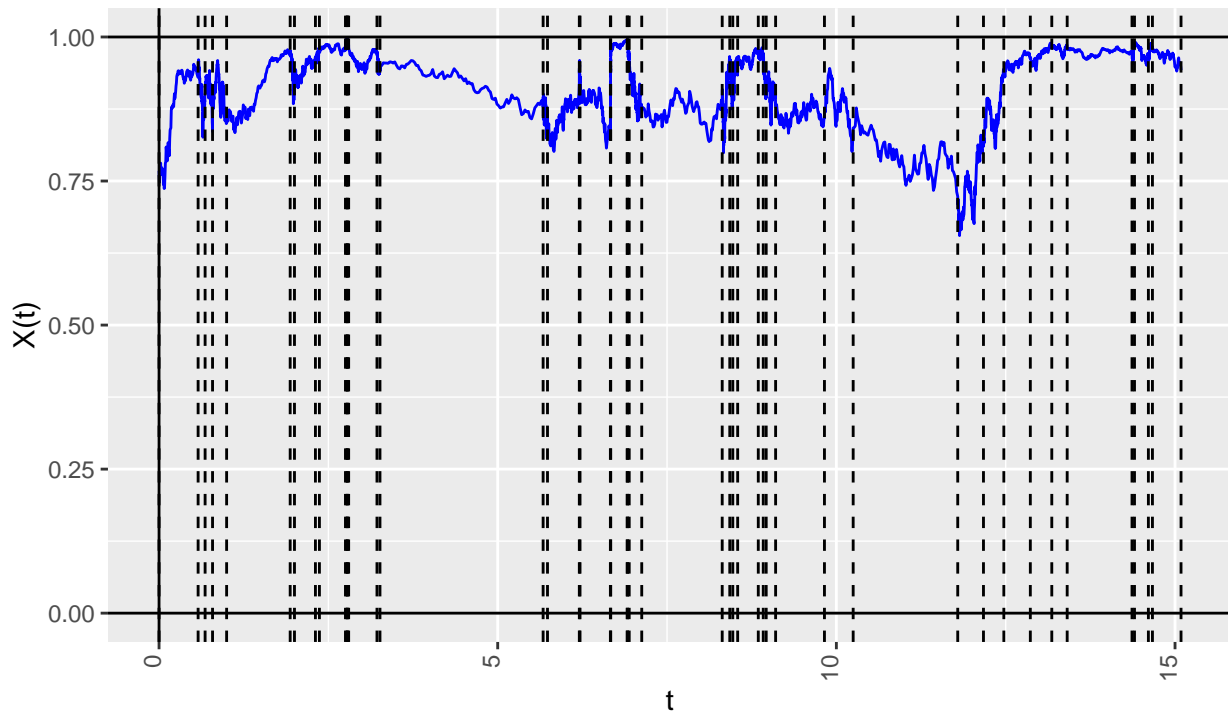


Figura 8.2 Simulación del proceso (220)-(221)

$\alpha=-0.4, \beta=1, k=1.5, 22$  cambios

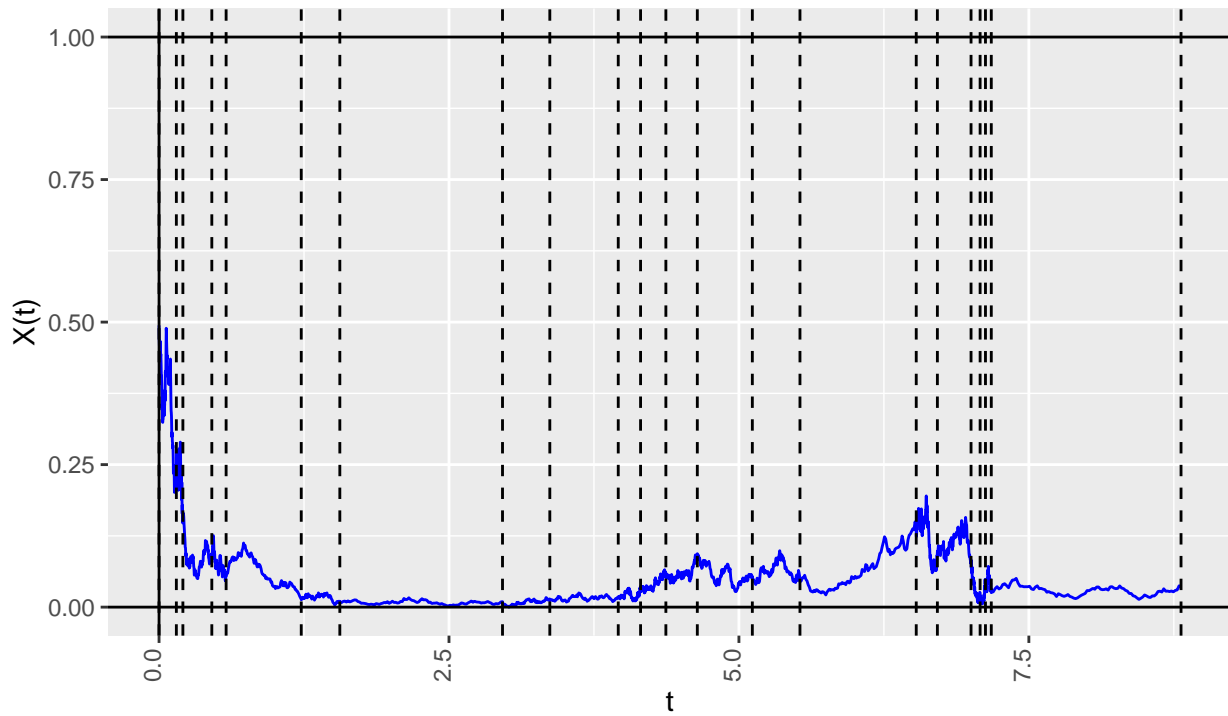


Figura 8.3 Simulación del proceso (220)–(221)

$\alpha=-0.4, \beta=-0.5, k=0.25, 22$  cambios

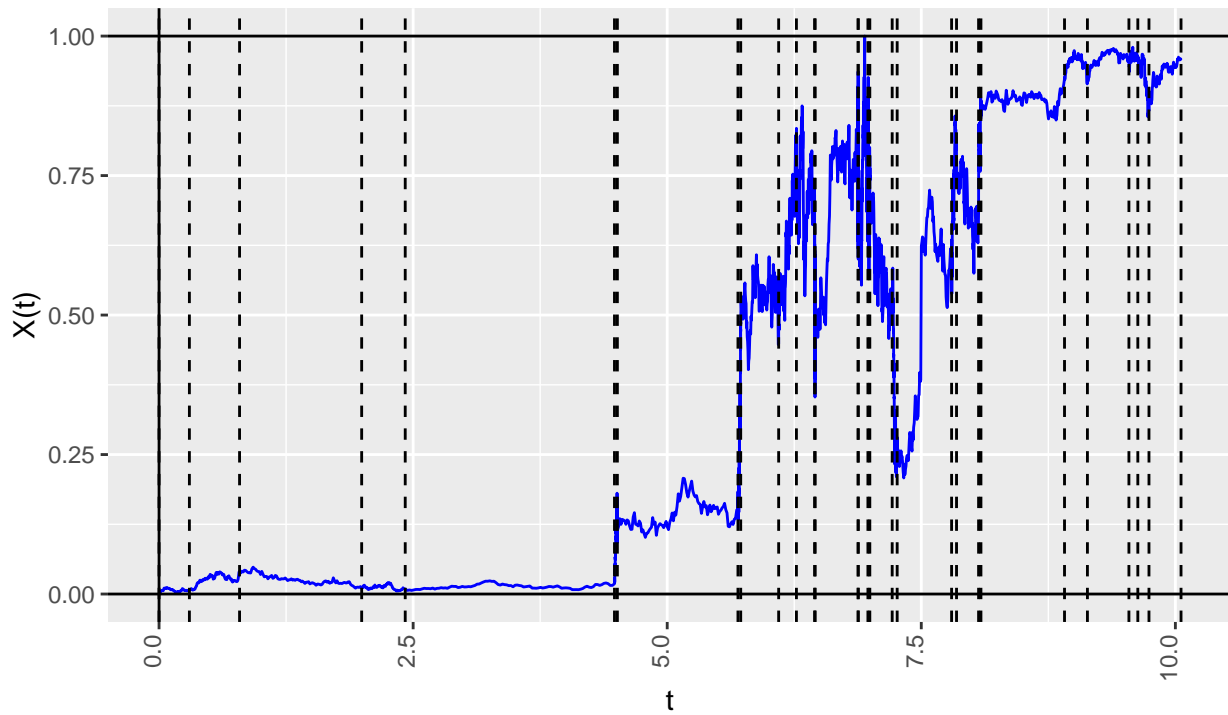


Figura 8.4 Simulación del proceso (220)–(221)

En [26], página 239 uno puede encontrarse la clasificación de los estados fronterizos 0 y 1 del modelo de Wright-Fisher que sólo considera efectos de mutación para las difusiones de Jacobi en términos de sus parámetros. Ambos estados 0 y 1 son fronteras *regulares* y/o *absorbentes* cuando  $-1 < \alpha, \beta < 0$ , respectivamente. Por otro lado, si  $\alpha, \beta \geq 0$ , entonces ambos estados 0 y 1 son fronteras *de entrada* y/o *reflectantes*. (Una frontera reflectante es aquella a la que no puede llegarse desde el interior del espacio de estados, aunque es posible considerar que el proceso inicie ahí. Una frontera absorbente es aquella de la que se puede entrar y salir).

Nótese que el estado de frontera 1 conserva este comportamiento en cualquier fase, pero la frontera 0 sólo conserva este comportamiento en la fase  $N$ . Para el resto de las fases  $i \in \{1, \dots, N-1\}$  la frontera 0 es reflectante, eso se debe a que  $\alpha + N + i > 0$ . Por lo tanto

$$\begin{cases} 1 \begin{cases} \text{es una frontera absorbente} & \text{para } -1 < \beta < 0, \\ \text{es una frontera reflectante} & \text{para } \beta \geq 0. \end{cases} \\ 0 \begin{cases} \text{es una frontera absorbente} & \text{para } -1 < \alpha < 0, i = N \\ \text{es una frontera reflectante} & \text{para } \alpha \geq 0, i = 1, \dots, N-1. \end{cases} \end{cases}$$

En las Figuras 8.1-8.4 puede observarse este comportamiento para cuatro situaciones distintas. En todas ellas se tienen  $N = 3$  fases. Dado que estamos enfocados en el comportamiento del proceso en las fronteras 0 y 1, no marcamos cuál es la fase en cada región. En la Figura 8.1 los parámetros fueron seleccionados de forma que las fronteras 0 y 1 sean reflectantes, es decir que la trayectoria nunca alcanzará los valores de frontera. En la Figura 8.2 los parámetros fueron seleccionados de forma que la frontera 0 sea reflectante y 1 sea absorbente. En la Figura 8.3 los parámetros fueron seleccionados de forma que la frontera 1 sea reflectante y 0 sea absorbente. En este caso, la única fase donde la trayectoria alcanza 0 es en la tercera. Para el resto de las fases, 0 es reflectante. Finalmente, en la Figura 8.4 los parámetros fueron seleccionados de forma que las fronteras 0 y 1 sean absorbentes. Este proceso puede alcanzar 1 en cualquier fase, pero sólo puede alcanzar 0 en la tercera fase.

De acuerdo con la **Definición 3.4.4**, el proceso es recurrente si y solo si  $\alpha, \beta \geq 0$ , y transitivo de otra forma. Lo anterior fue corroborado al analizar las soluciones numéricas de  $R_{x,y}$  por medio de la ecuación diferencial de segundo orden dada por (204). De manera similar, podemos asegurar ocupando la **Definición 3.4.7** que el proceso siempre es positivo recurrente para  $\alpha, \beta \geq 0$ , al analizar las soluciones numéricas de  $V(x)$  por medio de la ecuación diferencial de segundo orden dada por (206) donde  $G(x)$  está dado por (207).

Analicemos ahora un par de aspectos del componente discreto  $Y(t)$ . Primero los *tiempos de espera* de cada fase en función de los parámetros y luego veremos la *tendencia* de moverse hacia adelante o hacia atrás entre fases. Todo esto depende de la posición de  $X(t)$ .

Para los tiempos de espera, nótese que los elementos de la diagonal de  $Q(x)$  son

$$\begin{cases} Q_{11}(x) = -\frac{1}{1-x}(N-1)(\beta-k+1) & \text{para } i = 1, \\ Q_{ii}(x) = -\frac{1}{1-x}[(N-i)(\beta-k+i) + x(i-1)(N-i+k)] & \text{para } i = 2, \dots, N-1, \\ Q_{NN}(x) = -\frac{x}{1-x}(N-1)k & \text{para } i = N. \end{cases}$$

Por lo que si  $x$  está cerca de 1, los elementos de la diagonal son bastante grandes, eso quiere decir que las fases son *instantáneas*, por lo que los tiempos de espera en cada fase son bastante cortos hasta que  $x$  esté lejos de 1. Esto puede corroborarse en las Figuras 8.1-8.4. Cuando  $x$  se acerque a 0, se tiene que  $Q_{NN}(x)$  es bastante pequeño, por lo que la fase  $N$  se vuelve *absorbente* (por ende, si el proceso entra en esta fase y el valor de  $X(t)$  es cercano a 0, entonces el proceso tiende a pasar largos periodos de tiempo cerca de 0). Las fases intermedias (de 2 a  $N-1$ ) nunca son absorbentes, dado que  $Q_{ii}(x), i = 2, \dots, N-1$  nunca se acercan a 0. Por último, si  $k$  se acerca a  $\beta + 1$  entonces la fase 1 es absorbente. Para el resto de casos, el valor de las entradas diagonales y la posición de  $X(t)$  determinan el tiempo de espera en esa fase. Lo anteriormente mencionado puede visualizarse en las siguientes dos figuras.

$\alpha=1, \beta=2, k=0.2, 16$  cambios

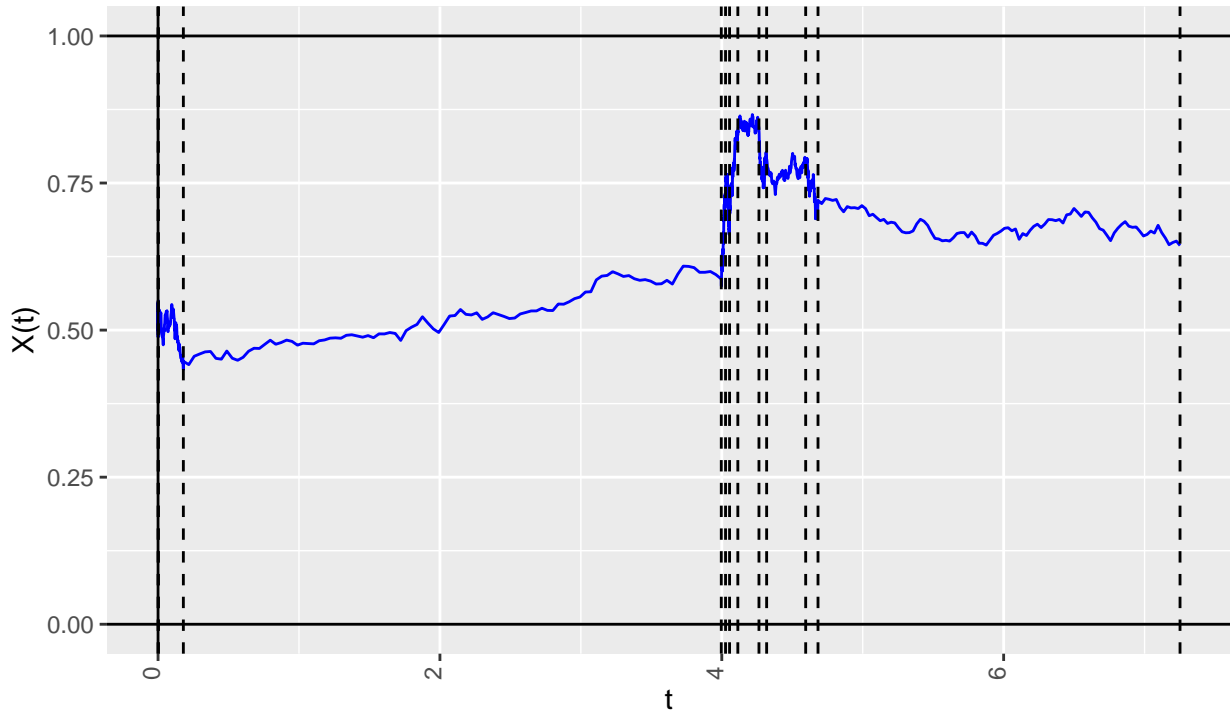


Figura 9.1 Simulación del proceso (220)–(221)

$\alpha=0.5, \beta=5, k=2.8, 18$  cambios

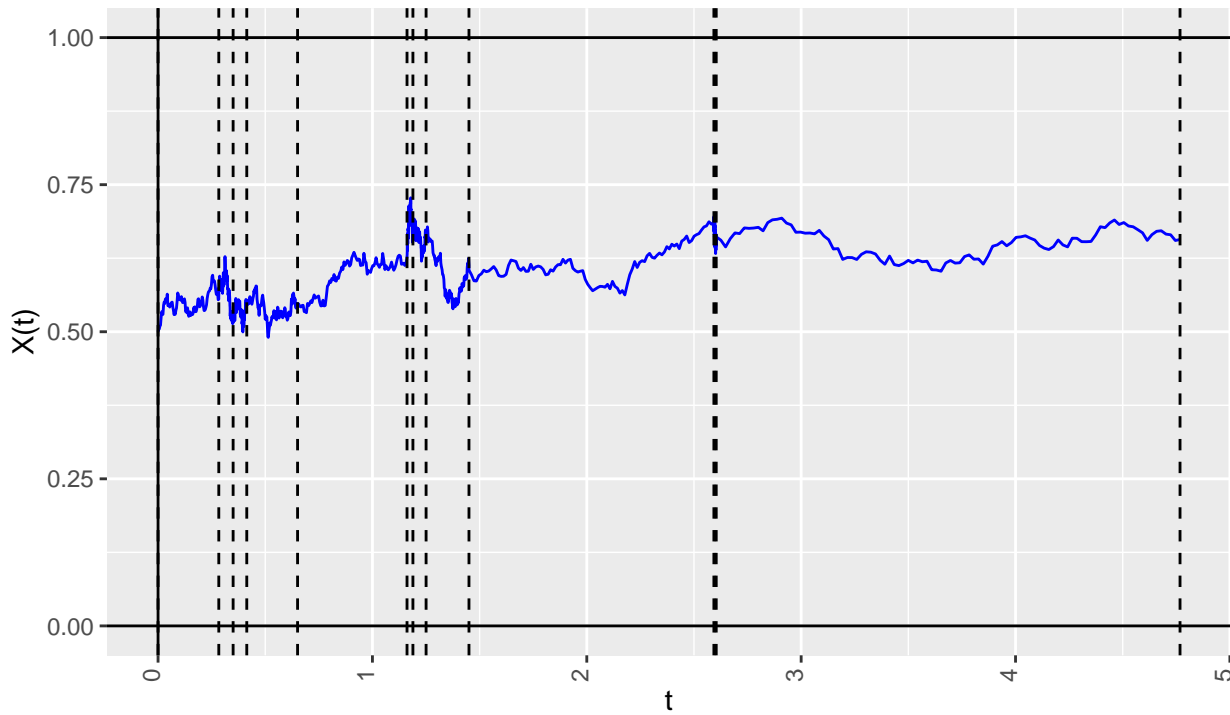


Figura 9.2 Simulación del proceso (220)–(221)

Ahora para analizar la tendencia, dado que el proceso se encuentra en la fase  $Y(t) = i$ , debemos revisar cuál es la probabilidad de moverse a la anterior o siguiente fase. Obviamente, cuando el proceso se encuentra en alguna de las fases fronterizas 1 o  $N$ , sólo puede moverse hacia adelante o hacia atrás, respectivamente. Estas probabilidades están dadas por

$$\begin{aligned}\mathbb{P}[Y(t) = i \rightarrow Y(t) = i + 1] &= \frac{\lambda_i(x)}{\lambda_i(x) + \mu_i(x)}, \\ \mathbb{P}[Y(t) = i \rightarrow Y(t) = i - 1] &= \frac{\mu_i(x)}{\lambda_i(x) + \mu_i(x)},\end{aligned}$$

donde  $x$  es el estado  $X(t)$  del proceso a tiempo  $t$  de transición de fases, además  $\lambda_i(x)$  y  $\mu_i(x)$  están dados por (216). Las tendencias hacia adelante o hacia atrás dependen de que  $\mu_i(x) < \lambda_i(x)$  o  $\mu_i(x) > \lambda_i(x)$ , respectivamente. Definimos los valores límite

$$x_0(i) = \frac{(N-i)(i+\beta-k)}{(i-1)(N-i+k)}, \quad i = 2, \dots, N-1. \quad (222)$$

Estos valores son las soluciones  $x$ -dependientes de la ecuación algebraica  $\lambda_i(x) = \mu_i(x)$ . Por ende, al momento del cambio de fase de  $Y(t) = i$  a la anterior o a la siguiente, tendremos una tendencia hacia adelante cuando  $x < x_0(i)$ , y una tendencia hacia atrás cuando  $x > x_0(i)$ . De manera que sea relevante, este valor límite debe ser menor a 1. De otra forma, siempre habrá una tendencia hacia adelante.

Estudiemos los casos donde el valor límite es significativo. Si  $x_0(i) > 1$  para todo  $i = 1, \dots, N-1$ , entonces siempre habrá tendencia hacia adelante. Esto sucede cuando

$$k < \frac{\beta+1}{N-1}.$$

De todas formas, si

$$k > \frac{(\beta+1)(N-i)}{N-1}, \quad (223)$$

para alguna  $i$ , puede haber tendencia hacia adelante o hacia atrás en la misma prueba, lo cual dependerá de la posición de  $X(t)$  en el momento de transición. Tenemos una máxima tendencia hacia atrás en (223) si  $i = 2$ . El rango de valores de  $k$  dependiendo de la fase  $Y(t) = i$  puede visualizarse en la Figura 10. Por lo tanto tenemos

$$\begin{cases} \text{Máxima tendencia hacia adelante} & \text{si } k < \frac{(\beta+1)}{N-1}, \\ \text{Máxima tendencia hacia atrás} & \text{si } k > \frac{(\beta+1)(N-2)}{N-1}, \\ \text{Tendencia hacia adelante/atrás} & \text{en otro caso.} \end{cases}$$

Esto explica por qué las fases 1 o  $N$  son absorbentes en términos de los valores de  $k$ .

### Estudio de tendencias: alpha=0.5,beta=1,N=5

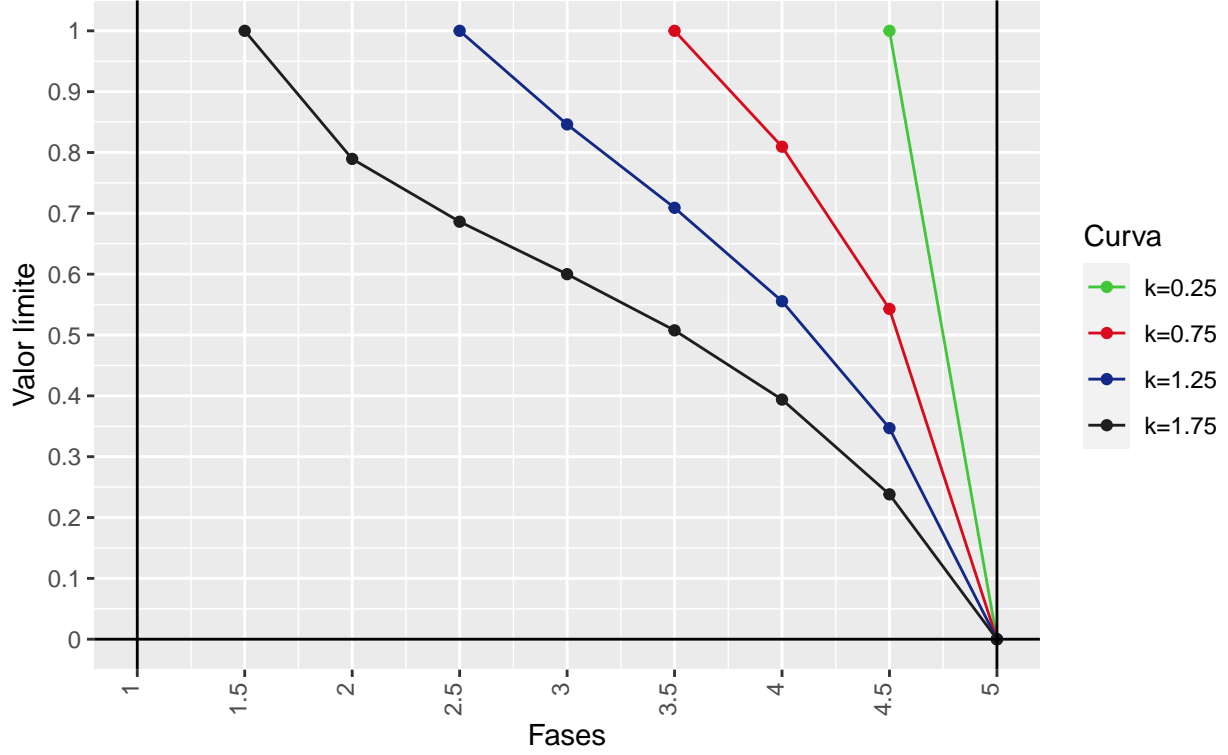


Figura 10: Valores límite para distintos valores de k

Finalmente, daremos una expresión explícita de la distribución invariante en forma vectorial  $\psi(y)$  (véase el **Teorema 3.4.9**). Podemos obtener una fórmula explícita en este caso dado que  $\mathcal{L}$  (213) es auto-adjunto y también tenemos una expresión explícita de la medida ortogonal  $W$ . Precisamente, el **Teorema 3.4.9** nos dice que:

$$\psi(y) = \left( \int_0^1 e_N^T W(y) e_N dy \right)^{-1} e_N^T W(y), \quad (224)$$

donde  $W$  está dado por (212). Dada la definición de  $H$ , es fácil demostrar que para todo  $j = 1, \dots, N$ ,

$$\psi_j(y) = y^{\alpha+N-j}(1-y)^\beta \binom{N-1}{j-1} \binom{\alpha+\beta+N}{\alpha} \frac{(\beta+N)(k)_{N-j}(\beta-k+1)_{j-1}}{(\alpha+\beta-k+2)_{N-1}}. \quad (225)$$

Nótese que esta distribución invariante (224) vectorial sólo es válida cuando el proceso es positivo recurrente, esto es, cuando  $\alpha, \beta \geq 0$ . Para el caso donde  $-1 < \alpha, \beta < 0$ , (224) también tiene sentido, pero ahora los puntos de la frontera serán absorbentes, por lo que la expresión correcta de la distribución invariante para este caso debe incluir los saltos o átomos en las fronteras 0 y 1 además de una densidad de la forma (224).

Para ejemplificar hemos graficado  $\psi_j(y), j = 1, \dots, N$  en (225) para  $N = 2, 3, 4, 5$ . Hemos fijado los parámetros  $\alpha = \beta = 1, k = \frac{5}{4}$ . Cabe resaltar que  $\int_0^1 \psi_j(y) dy < 1$  en cada componente, pero  $\sum_{j=1}^N \int_0^1 \psi_j(y) dy = 1$ .

Distribuciones invariantes:  $\alpha=1, \beta=1, k=1.25, N=2$

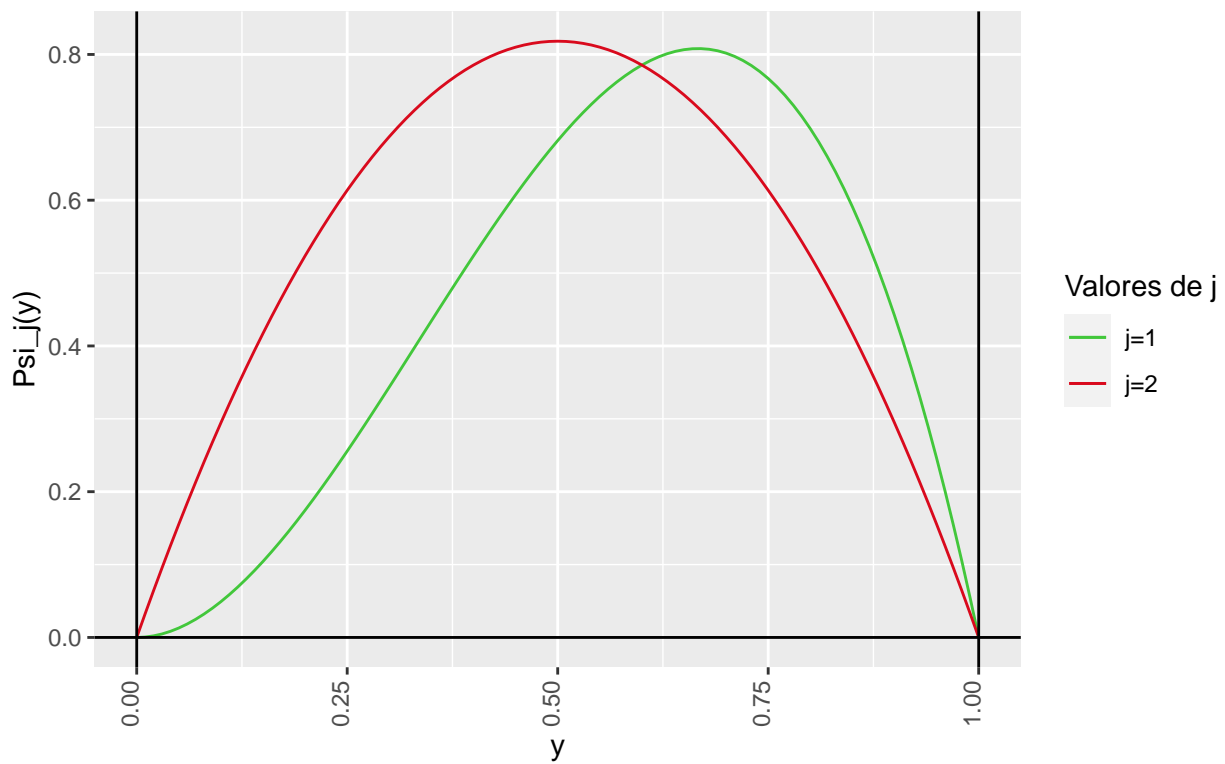


Figura 11.1: Componentes de distribución invariante,  $j=1,2$

Distribuciones invariantes:  $\alpha=1, \beta=1, k=1.25, N=3$

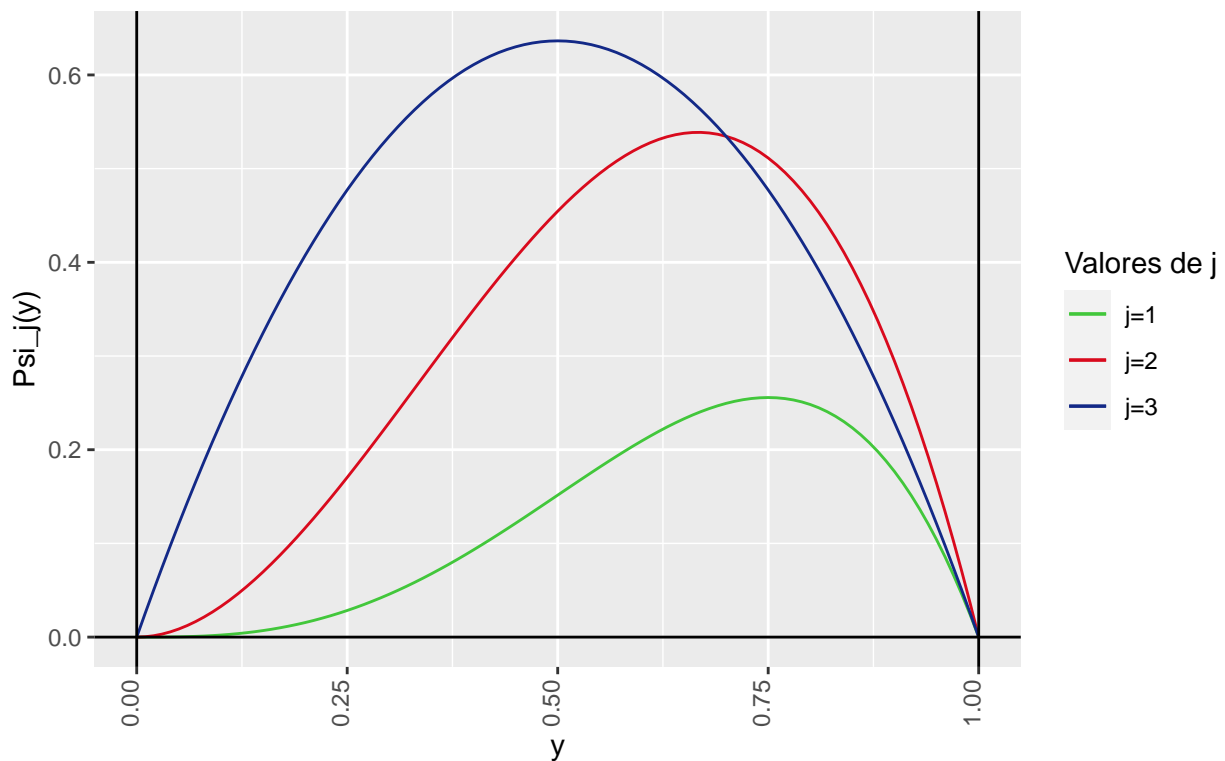


Figura 11.2: Componentes de distribución invariante,  $j=1,2,3$



Distribuciones invariantes:  $\alpha=1, \beta=1, k=1.25, N=4$

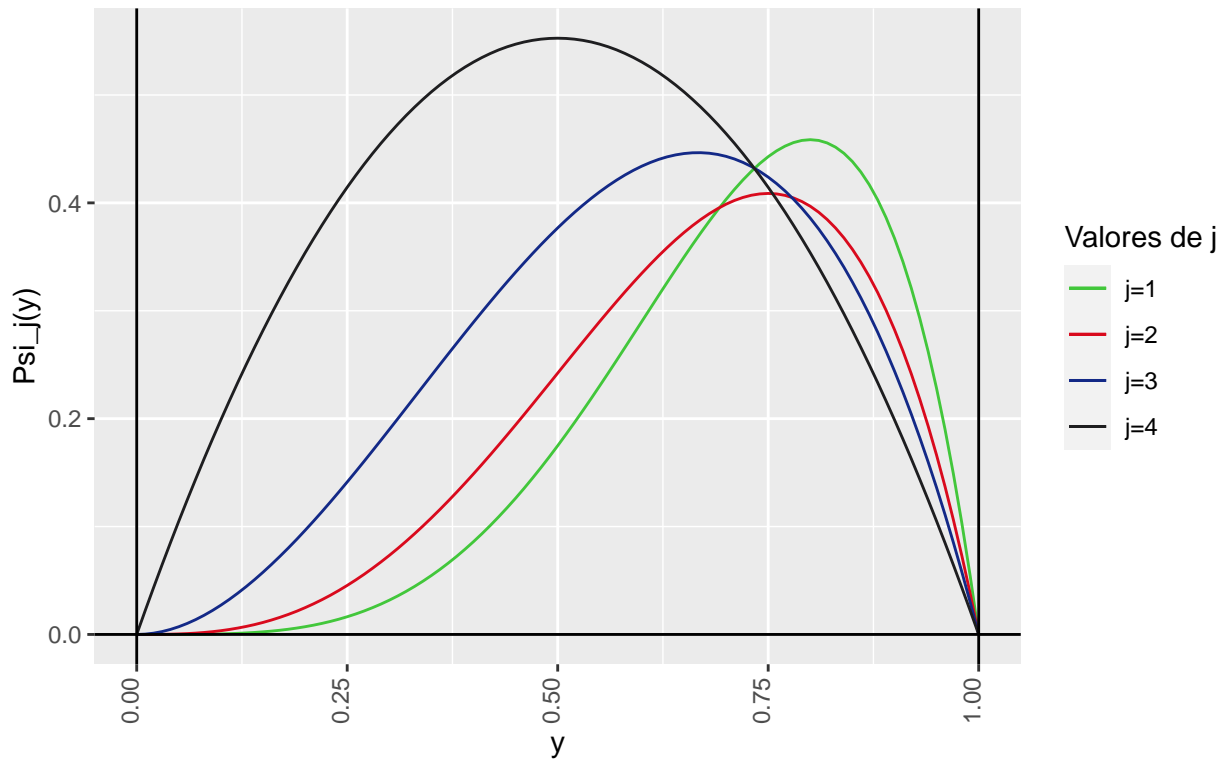


Figura 11.3: Componentes de distribución invariante,  $j=1,2,3,4$

Distribuciones invariantes:  $\alpha=1, \beta=1, k=1.25, N=5$

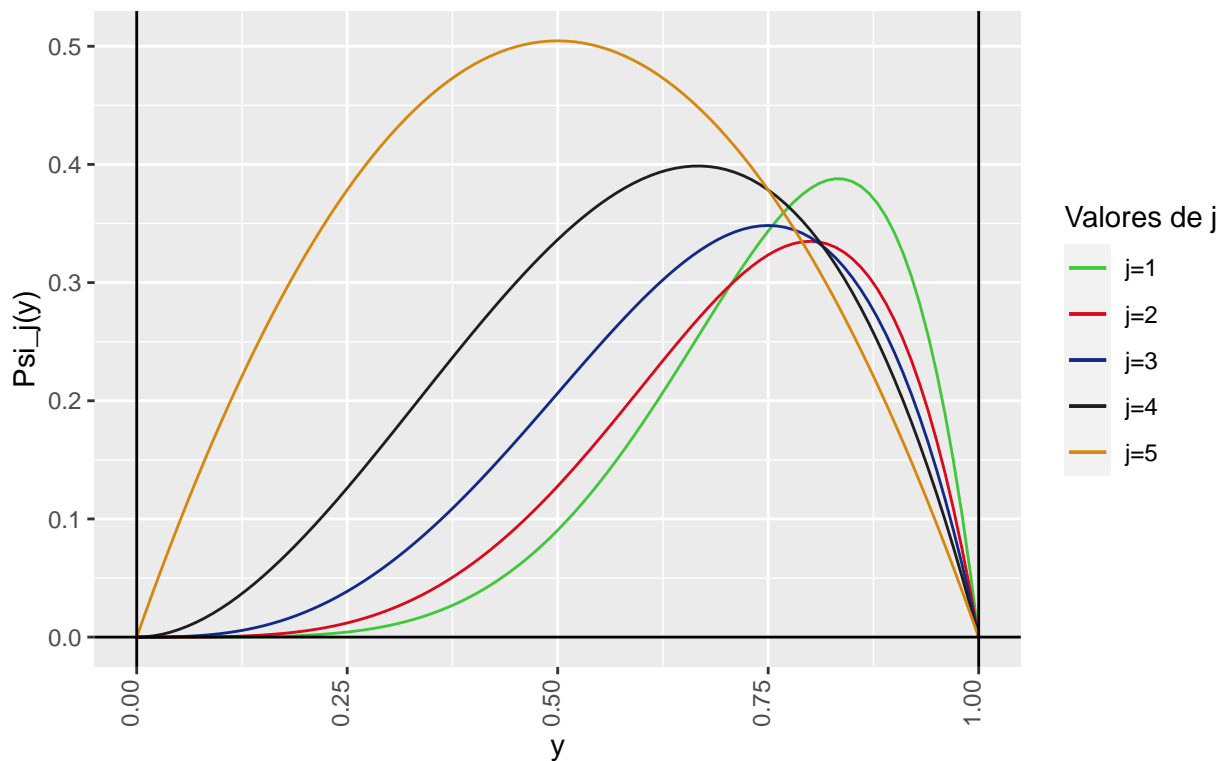


Figura 11.4: Componentes de distribución invariante,  $j=1,2,3,4,5$

De las Figuras 11.1-11.4 puede observarse que, para tiempos largos, si el proceso se encuentra en algún

estado en  $(0, \frac{1}{2})$ , entonces existe una probabilidad alta de que el proceso se localice en alguna de las últimas fases. Sin embargo, si el proceso está en algún estado en  $(\frac{1}{2}, 1)$ , entonces se tiene más o menos la misma probabilidad de que el proceso se ubique en cualquier fase (mayor la probabilidad para las primeras fases).

Estas interpretaciones pueden cambiar dependiendo del valor de  $k$ . Si  $k$  está cerca de  $\beta+1$  entonces tendremos la tendencia máxima hacia atrás y es más probable que el proceso esté en alguna de las fases iniciales. Sin embargo, si  $k$  está cerca de 0, entonces tendremos la tendencia máxima hacia adelante y es más probable que el proceso esté en alguna de las fases finales.

## Referencias

1. Anderson, W.J.: Continuous-time Markov chains: An applications-oriented approach. Springer Science & Business Media (2012).
2. Andrews, G.E. et al.: Special functions. Cambridge University Press (1999).
3. Arapostathis, A. et al.: Harnack's inequality for cooperative weakly coupled elliptic systems: Harnack's inequality. *Communications in partial differential equations*. 24, 9-10, (1999).
4. Berman, S.M.: A bivariate Markov process with diffusion and discrete components. *Stochastic Models*. 10, 2, 271–308 (1994).
5. Bhattacharya, R. et al.: Criteria for recurrence and existence of invariant measures for multidimensional diffusions. *The Annals of Probability*. 6, 4, (1978).
6. Bhattacharya, R.N., Waymire, E.C.: Stochastic processes with applications. SIAM (2009).
7. Björk, T.: Finite dimensional optimal filters for a class of Itô-processes with jumping parameters. *Stochastics: an international journal of probability and stochastic processes*. 4, 2, 167–183 (1980).
8. Blackwell, P.: Bayesian inference for Markov processes with diffusion and discrete components. *Biometrika*. 90, 3, 613–627 (2003).
9. Burkholder, D.L. et al.: Integral inequalities for convex functions of operators on martingales. En: *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* pp. 223–240 (1972).
10. Chen, Z.-Q., Zhao, Z.: Harnack principle for weakly coupled elliptic systems. *Journal of differential equations*. 139, 2, 261–282 (1997).
11. De la Iglesia, M.D.: Orthogonal polynomials in the spectral analysis of Markov processes. Birth death models and diffusion. *Encyclopedia of Mathematics and its Applications*, Cambridge University Press. 181, (2021).
12. De la Iglesia, M.D.: Spectral methods for bivariate Markov processes with diffusion and discrete components and a variant of the Wright–Fisher model. *Journal of Mathematical Analysis and Applications*. 393, 1, 239–255 (2012).
13. Durán, A.J., Grünbaum, F.A.: Orthogonal matrix polynomials satisfying second-order differential equations. *International Mathematics Research Notices*. 2004, 10, 461–484 (2004).
14. Dynkin, E.B.: Markov processes. En: *Markov processes*. pp. 77–104 Springer (1965).
15. Eizenberg, A., Freidlin, M.: Averaging principle for perturbed random evolution equations and corresponding Dirichlet problems. *Probability theory and related fields*. 94, 3, 335–374 (1993).
16. Eizenberg, A., Freidlin, M.: On the Dirichlet problem for a class of second order PDE systems with small parameter. *Stochastics and Stochastic Reports*. 33, 3-4, 111–148 (1990).
17. Friedman, A.: Stochastic differential equations and applications. En: *Stochastic differential equations*. pp. 75–148 Springer (2010).
18. Ghosh, M.K. et al.: Ergodic control of switching diffusions. *SIAM Journal on Control and Optimization*. 35, 6, 1952–1988 (1997).
19. Gilbarg, D., Trudinger, N.S.: Elliptic partial differential equations of second order. Springer (2015).
20. Golub, G.H. et al.: Matrix Computations. (1996).
21. Gronwall, T.H.: Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Annals of Mathematics*. 292–296 (1919).
22. Grünbaum, F.A., De la Iglesia, D.M.: Matrix valued orthogonal polynomials arising from group representation theory and a family of quasi-birth-and-death processes. *SIAM Journal on Matrix Analysis and Applications*. 30, 2, 741–761 (2008).

23. Guo, Q. et al.: The truncated Milstein method for stochastic differential equations with commutative noise. *Journal of Computational and Applied Mathematics*. 338, 298–310 (2018).
24. Hernández-Lerma, O., Sasserre, J.B.: Fatou’s lemma and Lebesgue’s convergence theorem for measures. *International Journal of Stochastic Analysis*. 13, 2, 137–146 (2000).
25. Il’in, A. et al.: Asymptotic expansions of solutions of integro-differential equations for transition densities of singularly perturbed switching diffusions: Rapid switchings. *Journal of mathematical analysis and applications*. 238, 2, 516–539 (1999).
26. Karlin, S., Taylor, H.E.: *A second course in stochastic processes*. Elsevier (1981).
27. Khasminskii, R.: *Stochastic stability of differential equations*. Springer Science & Business Media (2011).
28. Mao, X., Yuan, C.: *Stochastic differential equations with Markovian switching*. Imperial College Press (2006).
29. Protter, M.H., Weinberger, H.F.: *Maximum principles in differential equations*. Springer Science & Business Media (2012).
30. Royden, H.L., Fitzpatrick, P.: *Real Analysis*. Macmillan New York (1988).
31. Saw, J.G. et al.: Chebyshev inequality with estimated mean and variance. *The American Statistician*. 38, 2, 130–132 (1984).
32. Stroock, D.W., Varadhan, S.S.: *Multidimensional diffusion processes*. Springer (2007).
33. Wloka, J.T. et al.: *Boundary value problems for elliptic systems*. Cambridge University Press (1995).
34. Yin, G. et al.: Approximation methods for hybrid diffusion systems with state-dependent switching processes: numerical algorithms and existence and uniqueness of solutions. *SIAM Journal on Mathematical Analysis*. 41, 6, 2335–2352 (2010).
35. Yin, G.G., Zhang, Q.: *Continuous-time Markov chains and applications: a singular perturbation approach*. Springer (2012).
36. Yin, G.G., Zhu, C.: *Hybrid switching diffusions: properties and applications*. Springer Science & Business Media (2009).