



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

ERGODICIDAD EXPONENCIAL PARA PROCESOS DE LA CLASE AIMD

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL MATEMÁTICO

NIKOLAS ESTEBAN TAPIA MUÑOZ

PROFESOR GUÍA:
JOAQUÍN FONTBONA TORRES

MIEMBROS DE LA COMISIÓN:
JAIME SAN MARTÍN ARISTEGUI
DANIEL REMENIK ZISIS

SANTIAGO DE CHILE
2014

RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE: Ingeniero Civil Matemático
POR: Nikolas Esteban Tapia Muñoz
FECHA: 03/01/2014
PROFESOR GUÍA: Joaquín Fontbona Torres

ERGODICIDAD EXPONENCIAL PARA PROCESOS DE LA CLASE AIMD

En la presente memoria se estudia la convergencia al equilibrio de los procesos estocásticos pertenecientes a la clase Incremento Aditivo Decremento Multiplicativo (AIMD, por sus siglas en inglés), enmarcada en el contexto de los procesos de Markov deterministas por pedazos, introducidos por M.H.A. Davis en 1984.

En primer lugar se estudian los tiempos de entrada y salida de compactos de la forma $[0, x_0]$, primero para el caso con intensidad constante y luego en el caso general. Luego, mediante la construcción de un coupling específico, se usan estos resultados para establecer la recurrencia de cierto compacto, el que depende los datos del problema. También se prueba que el tiempo de entrada a este compacto tiene momento exponencial finito de algún orden. Como consecuencia, se establece la existencia de una medida invariante y se obtiene una tasa explícita de convergencia exponencial al equilibrio. Por último, se aplica el Teorema probado al caso del proceso TCP y se compara con las tasas exhibidas por Bardet *et al.* en 2011.

A Jeannette, Jorge y Felipe.

Agradecimientos

En primer lugar, quiero agradecer a las tres personas más importantes de mi vida; mis padres, Jeannette y Jorge, y mi hermano Felipe. Todo lo que soy y lo que he logrado se lo debo a Uds. Gracias por su apoyo constante, las oportunidades y los momentos agradables que hemos compartido. Gracias a mis primos Álvaro y Camila, por acompañarme siempre y ser una fuente infinita de alegría. Sepan que los quiero mucho.

Quiero agradecer a mis amigos, en especial a Gianfranco, David, Avelio, Ignacio V., Felipe, Luis y José con quienes he compartido experiencias enriquecedoras, tanto en lo académico como en lo personal. A Jocelyn y Nicolás por apoyarme cuando lo necesité y recibirme siempre en su casa; demás está decir que las incontables horas que pasamos juntos son un tesoro invaluable. A todas las personas que conocí durante mi paso por la Facultad, pero en especial a Emilio y Valentina G., Gonzalo F. y Evelyn, Ignacio C. y Valentina T., Rodolfo, Camila, Roberto, Sebastián R. y Gonzalo C., sin Uds. estos años no habrían pasado tan rápido. Por último, pero no menos importante, a mis compañeros de generación y en particular a Felipe, Sylvia, Sebastián B., Ricardo, Bastián y Víctor.

Gracias a los docentes del Departamento por la formación que me han entregado. A Rafael Correa por acercarme a las matemáticas y confiar siempre en mí. A Jaime San Martín, Servet Martínez y Daniel Remenik por inculcar en mí el amor hacia las Probabilidades, y por sobre todo a mi profesor guía, Joaquín Fontbona, por ayudarme siempre con nuevas ideas y por su preocupación tanto hacia mi trabajo como hacia mi persona; sin su ayuda nada de esto hubiera sido posible.

Finalmente, dar las gracias a todos los funcionarios del DIM por su atención y disponibilidad. Sin su presencia este departamento no sería capaz de funcionar como lo hace. Gracias por hacer más ameno el trabajo diario.

Tabla de contenido

Introducción	1
1. Preliminares	3
1.1. Semigrupos de operadores	4
1.1.1. Operadores en espacios de Funciones	5
1.2. Procesos de Markov	7
1.2.1. Problema de Martingala	9
1.3. Coupling	10
1.3.1. Coupling Maximal	11
1.3.2. Coupling entre procesos	12
1.4. Procesos de Markov deterministas por pedazos	13
1.5. La clase Incremento Aditivo Decremento Multiplicativo	15
1.5.1. Algunos ejemplos	16
2. Resultados	21
2.1. Tiempos de retorno	21
2.1.1. El caso de tasa constante	21
2.1.2. Caso general	25
2.2. Minoración de algunas probabilidades	28
2.3. Resultado Principal	30
2.4. Aplicación	37
Conclusión	39
Bibliografía	41

Índice de ilustraciones

1.1.	Trayectoria de la edad actual asociada a un proceso de renovación, con $F(x) = 1 - e^{-x}$ y $\mathcal{L}(X_0) = \mathcal{E}(1)$	18
1.2.	Trayectoria del proceso TCP con $X_0 \equiv 1$	19
2.1.	Trayectoria exitosa del coupling (\hat{X}^x, \hat{X}^y) con $x = 6, y = 2$	36
2.2.	Comportamiento de la solución $\tilde{\alpha}$ respecto de $p = 1 + \zeta$	38

Introducción

Los procesos de Markov deterministas por pedazos (PDMP por sus siglas en inglés) fueron introducidos en 1984 por M.H.A. Davis en su paper seminal *Piecewise-Deterministic Markov Process: A General Class of Non-Diffusion Stochastic Models* [7]. La razón principal de Davis para introducir esta clase de procesos fue que presentan una alternativa a los modelos de difusión clásicos, esto es, la dinámica ya no es gobernada por un movimiento Browniano, sino que por una ecuación diferencial ordinaria, y la única aleatoriedad se encuentra en los saltos.

Recientemente se han estudiado modelos de este tipo, en donde el espacio de estados es \mathbb{R}_+ , la dinámica determinista es lineal y la mecánica de saltos es multiplicativa, esto es, la posición después de cada salto es escogida multiplicando el estado actual por una v.a. con una distribución dada, concentrada en el intervalo $[0, 1)$. Es por esto que esta subclase es llamada Incremento Aditivo Decremento Multiplicativo (AIMD por sus siglas en inglés) y es en ésta donde se enfoca este trabajo.

El ejemplo más emblemático de proceso perteneciente a la clase AIMD es el llamado proceso TCP, que modela el tamaño de una ventana de transmisión de datos en el protocolo de internet del mismo nombre. Dicho modelo aparece como el límite en escalamiento de un algoritmo discreto de tipo Markoviano para regular la congestión en un canal de transmisión de datos. Varias propiedades de este límite se estudian en [10, 18].

A nivel general, una pregunta de interés en el contexto de procesos de Markov (de la clase PDMP o no) es en primer lugar, si poseen medidas invariantes o estacionarias. Esta interrogante está bastante estudiada y da origen a la teoría de las cadenas de Harris [1, Sección VII.3]. Una vez que se ha establecido la existencia de estas medidas, cabe preguntarse por la velocidad de convergencia de la ley a tiempo t del proceso subyacente a esta medida estacionaria en alguna distancia adecuada. Existen técnicas generales aplicables a cualquier proceso de Markov a tiempo continuo para deducir esta convergencia en varios casos. Entre estas destacan las expuestas por Dufour & Costa [9], Lindvall [17], Roberts & Rosenthal [20], y Thórisson [21].

En el caso del TCP, técnicas particulares de *coupling* han sido utilizadas para deducir tasas de convergencia exponencial a la medida invariante en la distancia de Wasserstein W_p y en variación total. Por ejemplo, se puede consultar los trabajos de Bardet *et al.* [4] y Chafaï *et al.* [6]. En Azaïs *et al.* [2] se presenta un resumen de los últimos resultados obtenidos para este proceso. La técnica desarrollada en este trabajo está inspirada en las utilizadas allí (en particular [4]).

Otros ejemplos de importancia están relacionados con la teoría de Procesos de Renovación. En este caso se incluyen la vida corriente y restante asociadas a un proceso de renovación. Las

técnicas utilizadas en este contexto difieren de las usadas en el caso del proceso TCP en que no permiten encontrar tasas explícitas de convergencia al equilibrio. En este sentido, aquí se presenta un resultado unificador pues es aplicable a ambos casos (y a muchos otros) y permite obtener tasas de convergencia explícitas, aunque más complicadas que las ya existentes. A groso modo, se probará el siguiente resultado

Teorema *Dadas dos leyes iniciales μ y ν con primer momento finito, existen $\hat{\gamma} > 0$ y una constante $C = C(\mu, \nu, \gamma) > 0$ tal que*

$$\|\mathbb{P}_\mu(X_t \in \cdot) - \mathbb{P}_\nu(X_t \in \cdot)\|_{VT} \leq C e^{-\gamma t}$$

para todo $\gamma < \hat{\gamma}$ y $t > 0$.

La presente memoria se organiza como sigue. En el Capítulo 1 se presentan las herramientas básicas para el correcto desarrollo de los resultados posteriores. En particular, se presenta una breve introducción a la Teoría de Semigrupos de Operadores (Sección 1.1), a los procesos de Markov (Sección 1.2), a la técnica de coupling (Sección 1.3) y a las clases PDMP (Sección 1.4) y AIMD (Sección 1.5).

Seguidamente, en el Capítulo 2 se presentan los resultados principales de este trabajo. En primer lugar, en la Sección 2.1 se estudian los tiempos de retornos a ciertos compactos, primero para el caso con intensidad constante y luego en el caso general. Después, en la Sección 2.2 se prueba una serie de lemas técnicos relacionados con los couplings que se construirán más adelante. Finalmente, en la Sección 2.3 se enuncia y prueba el resultado principal de la memoria.

Capítulo 1

Preliminares

En este capítulo se presenta el marco teórico bajo el que se desarrolla este trabajo. En su mayoría, los resultados aquí presentados se dan sin demostración pues corresponden a teoremas clásicos en la literatura y el contexto en el que aparecen es más general de lo necesario. Aún así, emergen como resultado de la revisión bibliográfica y como complemento a los conocimientos adquiridos durante la carrera.

Dado un espacio medible (E, \mathcal{E}) , se denota por $\mathcal{P}(E)$ a la colección de las medidas de probabilidad en \mathcal{E} y $B(E)$ será el espacio vectorial de las funciones medibles y acotadas dotado de la norma uniforme. Si E es un espacio topológico localmente compacto, siempre se considera su σ -álgebra Boreliana, $C_b(E) \subset B(E)$ será el espacio de las funciones continuas y acotadas, y $C_0(E) \subset C_b(E)$ el espacio de las funciones continuas que se anulan en infinito; finalmente $C_c(E) \subset C_0(E)$ es el espacio de las funciones continuas a soporte compacto. Además, \mathcal{K} y \mathcal{O} son la colección de subconjuntos compactos y abiertos de E , respectivamente. Si X es un espacio de Banach, entenderemos por *funcional en X* a una función $T : X \rightarrow \mathbb{R}$ y por *operador en X* a una función $T : X \rightarrow X$.

Para un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, las funciones $X : \Omega \rightarrow E$ que son \mathcal{F} - \mathcal{E} -medibles son llamados elementos aleatorios. En el caso en que $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ se dice que X es una variable aleatoria (v.a.). Todo elemento aleatorio X induce una medida $\mathcal{L}(X) \in \mathcal{P}(E)$, llamada *ley de X* , mediante la relación

$$\mathcal{L}(X)(A) = \mathbb{P}(X^{-1}(A)) \doteq \mathbb{P}(X \in A), \quad A \in \mathcal{E}.$$

Una filtración en \mathcal{F} es una colección de sub- σ -álgebras $(\mathcal{F}_t)_{t \geq 0}$ tales que $\mathcal{F}_s \subset \mathcal{F}_t$ cada vez que $s < t$. Se define

$$\mathcal{F}_\infty = \sigma \left(\bigcup_{t \geq 0} \mathcal{F}_t \right).$$

Un proceso adaptado $(M_t)_{t \geq 0}$ tal que $M_t \in L^1$ para cada $t \geq 0$ se dice submartingala (resp. sobremartingala) si $\mathbb{E}(X_t \mid \mathcal{F}_s) \geq X_s$ (resp. $\mathbb{E}(X_t \mid \mathcal{F}_s) \leq X_s$) para todo par de tiempos $t > s$. Si M_t es un sub y sobremartingala a la vez, se dirá que es una martingala. La teoría de martingalas es relativamente moderna, rica en resultados, y tiene diversas aplicaciones en varias ramas de las probabilidades y las matemáticas. Se pueden encontrar los resultados principales y un desarrollo detallado de esta teoría y sus aplicaciones en [5, 16, 19].

1.1. SEMIGRUPOS DE OPERADORES

Definición 1.1 Un núcleo de transición en (E, \mathcal{E}) es una función $N : E \times \mathcal{E} \rightarrow \mathbb{R} \cup \{\infty\}$ tal que

1. La aplicación $x \mapsto N(x, A)$ es medible para cada $A \in \mathcal{E}$.
2. La aplicación $A \mapsto N(x, A)$ para todo $x \in E$.

Si además $N(x, \cdot) \in \mathcal{P}(E)$ para todo $x \in E$, se dice que N es una probabilidad de transición.

Dados dos núcleos de transición N y M , es posible definir un tercero mediante

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

Para $f : E \rightarrow \mathbb{R}$ medible y μ una medida en \mathcal{E} , se pueden considerar las acciones de N por la izquierda sobre f y por la derecha sobre μ ,

$$Nf(x) = \int_E f(y) N(x, dy) \tag{1.1}$$

y

$$\mu N(A) = \int_E N(x, A) \mu(dx). \tag{1.2}$$

Es claro que estas acciones dejan invariantes ambos espacios, esto es, Nf es nuevamente una función medible y μN es una medida.

Definición 1.2 Una familia $(P_t)_{t \geq 0}$ de probabilidades de transición se dice función de transición si $P_0(x, \cdot) = \delta_x$ para todo $x \in E$, y satisface la ecuación de Chapman–Kolmogorov

$$P_{t+s}(x, A) = \int_E P_s(y, A) P_t(x, dy) \tag{1.3}$$

1.1. Semigrupos de operadores

Una familia de operadores lineales acotados $(T_t)_{t \geq 0}$ en un espacio de Banach X se dice *semigrupo* si $T_0 = I$ y $T_{t+s} = T_t T_s$ para todo $t, s \geq 0$. Un semigrupo $(T_t)_{t \geq 0}$ se dice *fuertemente continuo* si $\|T_t f - f\| \rightarrow 0$ para todo elemento $f \in X$; diremos que es un semigrupo de contracciones si $\|T_t\| \leq 1$ para todo $t \geq 0$.

EJEMPLO 1.1. Una función de transición $(P_t)_{t \geq 0}$ define naturalmente una familia de operadores, la que llamaremos igualmente $(P_t)_{t \geq 0}$, mediante su acción por la izquierda (ecuación (1.1)) sobre diversos espacios, por ejemplo, sobre $B(E)$, o sobre $C_b(E)$ o $C_0(E)$ si E es un espacio topológico localmente compacto. Este semigrupo no necesariamente debe ser fuertemente continuo.

Usando solamente la definición es posible demostrar ciertas propiedades de regularidad de estos semigrupos.

Proposición 1.3 (Corolario 1.1.2 en [11]) Si $(T_t)_{t \geq 0}$ es un semigrupo fuertemente continuo, entonces para todo $f \in X$ la aplicación $t \mapsto T_t f$ es continua de $[0, \infty)$ en X .

El *generador infinitesimal* (o simplemente generador) del semigrupo $(T_t)_{t \geq 0}$ es el operador lineal (posiblemente no acotado) A en X definido por

$$Af = \lim_{t \downarrow 0} \frac{T_t f - f}{t},$$

cuyo dominio $\mathcal{D}(A)$ es el subespacio de todos los vectores $f \in X$ tales que este límite existe.

Proposición 1.4 (Proposición 1.1.5 en [11]) *Sea $(T_t)_{t \geq 0}$ un semigrupo fuertemente continuo en X , con generador A .*

1. Para $f \in X$ y $t \geq 0$, se tiene que $\int_0^t T_s f \, ds \in \mathcal{D}(A)$ y

$$T_t f - f = A \int_0^t T_s f \, ds.$$

2. Si $f \in \mathcal{D}(A)$ y $t \geq 0$, entonces $T_t f \in \mathcal{D}(A)$ y

$$\frac{d}{dt} T_t f = A T_t f = T_t A f.$$

3. Si $f \in \mathcal{D}(A)$ y $t \geq 0$, entonces

$$T_t f - f = \int_0^t A T_s f \, ds = \int_0^t T_s A f \, ds.$$

Por último, es interesante saber cuándo un operador A genera un semigrupo fuertemente continuo. Una respuesta está dada por el Teorema de Hille–Yosida. Un operador A se dice *disipativo* si $\|\lambda f - A f\| \geq \lambda \|f\|$ para todo $f \in \mathcal{D}(A)$.

Teorema 1.5 (Hille & Yosida, Teorema 1.2.6 en [11]) *El operador lineal A en X es el generador de un semigrupo fuertemente continuo si y solamente si*

1. $\mathcal{D}(A)$ es denso en X .
2. A es disipativo.
3. $\text{Im}(\lambda - A) = X$ para algún $\lambda > 0$.

1.1.1. Operadores en espacios de Funciones

Durante esta sección E será un espacio topológico separado y localmente compacto. Un funcional T en $C_0(E)$ se dice *positivo* si $Tf \geq 0$ cada vez que $f \geq 0$. Una medida μ en $\mathcal{B}(E)$ se dice *regular interior* si

$$\mu(A) = \sup\{\mu(K) : K \in \mathcal{K}, K \subset A\}$$

para todo $A \in \mathcal{B}(E)$. La medida μ se dice *regular exterior* si

$$\mu(A) = \inf\{\mu(F) : F \in \mathcal{O}, A \subset F\}$$

1.1. SEMIGRUPOS DE OPERADORES

para todo $A \in \mathcal{B}(E)$. Por último, μ se dice *regular* si es regular interior y exterior simultáneamente.

Dada una medida finita y regular μ , es posible definir un funcional T_μ actuando en $B(E)$ mediante

$$T_\mu f = \int_E f(y) \mu(dy). \quad (1.4)$$

Es fácil ver que entonces T_μ es positivo, continuo y $\|T_\mu\| = \mu(E)$. Una pregunta natural entonces es saber si todos los funcionales en $B(E)$ tienen esta forma, esto es, interesa saber si dado T actuando sobre $B(E)$ es posible encontrar una medida finita y regular μ que represente a T en el sentido de la fórmula anterior. Lamentablemente esto no es así, como lo muestra el siguiente ejemplo.

EJEMPLO 1.2. Sea $E = \mathbb{R}$, y B' el subespacio de $B(E)$ constituido por las funciones f tales que $f(\infty) \doteq \lim_{x \rightarrow \infty} f(x)$ existe. En B' se define el funcional T mediante $Tf = f(\infty)$. Naturalmente T es continuo y positivo, y $\|T\| \leq 1$. Por el Teorema de Extensión de Hahn–Banach T tiene una (única) extensión continua y positiva a todo $B(E)$, la que seguiremos llamando T , y es tal que $\|T\| \leq 1$.

Veremos que no es posible que exista una medida μ finita y regular que represente a T . Si esto no fuera así, dado un compacto K y $\varepsilon > 0$ se puede encontrar una función $0 \leq f \leq 1$, tal que $f \equiv 1$ en K y $f \equiv 0$ en K_ε^c . Aquí, $K_\varepsilon = \{x : d(x, K) \leq \varepsilon\}$. Luego

$$0 \leq \mu(K) \leq \int_E f d\mu = Tf = f(\infty) = 0,$$

lo que junto con la regularidad de μ implica que $\mu \equiv 0$ lo que no puede ser.

El problema principal con el ejemplo anterior es que el espacio $B(E)$ es *muy grande*, es decir, contiene muchas funciones cuyo comportamiento no es adecuado para poder encontrar una representación como en (1.4). El siguiente resultado, llamado Teorema de Representación de Riesz–Markov, es de gran importancia y es una especie de recíproco del ejemplo anterior.

Teorema 1.6 (Teorema 7.17 en [12]) *Si T es funcional positivo y continuo en $C_0(E)$, entonces existe una única medida finita y regular μ sobre $\mathcal{B}(E)$ tal que*

$$Tf = \int_E f(y) \mu(dy).$$

Más aún, $\mu(E) = \|T\|$.

En este contexto, los semigrupos fuertemente continuos y positivos en $C_0(E)$ son llamados *semigrupos de Feller*. Si el semigrupo asociado a una función de transición es de Feller, se dice que esta es *de Feller*. Asociado a toda función de transición de Feller se tiene entonces un generador infinitesimal (el de su semigrupo), al que por extensión llamaremos generador de la función de transición.

Como consecuencia del Teorema de Riesz–Markov se tiene la siguiente

Proposición 1.7 (Proposición III.2.2. en [19]) *A todo semigrupo de Feller $(T_t)_{t \geq 0}$ es posible asociarle una única función de transición $(P_t)_{t \geq 0}$ tal que*

$$T_t f(x) = P_t f(x)$$

para toda función $f \in C_0(E)$ y $x \in E$.

1.2. Procesos de Markov

Sea S un espacio métrico, el que se supone completo y separable.

Definición 1.8 *Un proceso estocástico $(X_t)_{t \geq 0}$ con valores en S , definido en el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ se dice de Markov con respecto a la filtración $(\mathcal{F}_t)_{t \geq 0}$ si es adaptado y*

$$\mathbb{E}(f(X_{t+s}) \mid \mathcal{F}_s) = \mathbb{E}(f(X_{t+s}) \mid X_s)$$

para todo $t, s \geq 0$ y toda función $f \in B(S)$. La distribución inicial de X es la medida de probabilidad $\nu \in \mathcal{P}(S)$ definida por $\nu(A) = \mathbb{P}(X_0 \in A)$. Una función de transición $(P_t)_{t \geq 0}$ es una función de transición para X si

$$\mathbb{E}(f(X_{t+s}) \mid X_s) = P_t f(X_s) \quad (1.5)$$

para todo $t, s \geq 0$ y $f \in B(S)$.

En particular se tiene que $P_t(x, A) = \mathbb{P}(X_t \in A \mid X_0 = x)$ para todo $A \in \mathcal{B}(S)$, $x \in S$ y $t \geq 0$; esto es, $P_t(x, \cdot)$ es la ley de X_t condicional a que $X_0 = x$. Como consecuencia de la ecuación (1.5) se obtiene la llamada *propiedad de Markov*,

$$\mathbb{E}(f(X_{t+s}) \mid \mathcal{F}_s) = \mathbb{E}(f(X_t) \mid X_0) \quad (1.6)$$

Un primer ejemplo de proceso de Markov lo constituye el Movimiento Browniano.

EJEMPLO 1.3. Considerando $S = \mathbb{R}$ y el núcleo del calor

$$p_t(x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}},$$

el proceso de Markov B cuya función de transición está dada por

$$P_t(x, A) = \int_A p_t(x, y) dy$$

es llamado Movimiento Browniano. Si además $\nu = \delta_0$, se dice que B es un Movimiento Browniano estándar.

Un resultado clásico en la literatura es el siguiente, que caracteriza la distribución de un proceso de Markov en términos de su distribución inicial y función de transición.

Teorema 1.9 (Teorema III.1.5 en [19]) *Sea $(P_t)_{t \geq 0}$ una función de transición y $\nu \in \mathcal{P}(S)$. Existe un proceso de Markov X definido en el espacio de probabilidad $(S^{\mathbb{R}+}, \mathcal{B}(S^{\mathbb{R}+}), \mathbb{P}_\nu)$, tal que $(P_t)_{t \geq 0}$ es una función de distribución para X y ν es su distribución inicial.*

La realización en el espacio de probabilidad $(S^{\mathbb{R}+}, \mathcal{B}(S^{\mathbb{R}+}), \mathbb{P}_\nu)$ es llamada *realización canónica de X* . En esta realización, se tiene que $X_t(\omega) = \omega(t)$ es simplemente el operador de proyección en cada coordenada $t \geq 0$. Denotaremos \mathbb{P}_x a la medida de probabilidad correspondiente cuando $\nu = \delta_x$, para $x \in E$. La medida \mathbb{P}_x es, por definición, tal que $\mathbb{P}_x(X_0 = x) = 1$. Más aún, es posible probar que $\mathbb{P}_x(X_t \in A) = P_t(x, A)$ y en consecuencia se tiene la fórmula

$$P_t f(x) = \mathbb{E}_x(f(X_t)) \quad (1.7)$$

1.2. PROCESOS DE MARKOV

para toda $f \in B(S)$. Usando esta notación, la propiedad de Markov (1.6) se reescribe como

$$\mathbb{E}_x(f(X_{t+s}) \mid \mathcal{F}_s) = \mathbb{E}_{X_s}(f(X_t)). \quad (1.8)$$

Es importante notar que este Teorema no asegura ningún tipo de propiedades de regularidad de las trayectorias de X .

Resulta que es suficiente conocer \mathbb{P}_x para todo $x \in E$, pues el siguiente resultado indica que basta con *promediar* con respecto a la condición inicial estas cantidades.

Teorema 1.10 (Teorema III.1.6 en [19]) *Sea Z una v.a. \mathcal{F}_∞ medible y positiva (o acotada). Entonces la aplicación $x \mapsto \mathbb{E}_x(Z)$ es medible, y para cualquier $\nu \in \mathcal{P}(E)$ se tiene que*

$$\mathbb{E}_\nu(Z) = \int_E \mathbb{E}_x(Z) \nu(dx).$$

Tomando $Z = \mathbf{1}_A(X_t)$, con $A \in \mathcal{B}(S)$, se deduce que $\mathbb{P}_\nu(X_t \in A) = \nu P_t(x, A)$, es decir, la medida νP_t es la ley de X_t bajo \mathbb{P}_ν .

En realidad, la propiedad de Markov (1.8) puede ser escrita de forma un poco más general considerando la realización canónica de X . Si definimos, para $\omega \in S^{\mathbb{R}^+}$ y $t \geq 0$, el operador de traslación $\theta_t \omega(s) = \omega(t+s)$ entonces, si Z es una v.a. \mathcal{F}_∞ medible y positiva se tiene que

$$\mathbb{E}_x(Z \circ \theta_t \mid \mathcal{F}_t) = \mathbb{E}_{X_t}(Z)$$

para todo $t \geq 0$ y $x \in S$.

Si la función de transición de un proceso de Markov X resulta ser de Feller, se dirá que X es un *proceso de Feller*. En este caso, X satisface la llamada *propiedad de Markov fuerte*: si $T \geq 0$ es un tiempo de parada finito c.s. y Z es una v.a. \mathcal{F}_∞ medible y positiva, entonces

$$\mathbb{E}_x(Z \circ \theta_T \mid \mathcal{F}_T) = \mathbb{E}_{X_T}(Z). \quad (1.9)$$

Si X es un proceso de Feller se llama, por extensión, generador infinitesimal de X al generador infinitesimal de su función de transición. De la definición, se puede observar que si $f \in B(S)$ entonces

$$\mathbb{E}_x(f(X_{t+h}) - f(X_t) \mid \mathcal{F}_t) = P_h f(X_t) - f(X_t)$$

y luego, si A es el generador de X y $f \in \mathcal{D}(A)$ se tiene que

$$\mathbb{E}_x(f(X_{t+h}) - f(X_t) \mid \mathcal{F}_t) = hAf(X_t) + o(h).$$

Esta identidad puede ser interpretada como una descripción infinitesimal de las transiciones de X . En conexión con la Proposición 1.7 se tiene el siguiente resultado.

Teorema 1.11 *Sea S localmente compacto y separable, y $(T_t)_{t \geq 0}$ un semigrupo de Feller en $C_0(S)$. Entonces, para cada $\nu \in \mathcal{P}(S)$ existe un proceso de Feller X con semigrupo $(T_t)_{t \geq 0}$ y distribución inicial ν . Más aún, X posee una modificación cádlág.*

Muchas veces es difícil obtener expresiones explícitas para la función de transición de un proceso de Markov. En general resulta más fácil describir el comportamiento *infinitesimal* de dicho proceso; sin embargo, esto no es siempre posible. Recordando que toda función de transición define un semigrupo de operadores, y recíprocamente en el caso de semigrupos de Feller, el Teorema 1.9 permite construir un proceso de Markov a través de su generador infinitesimal.

Se dice que un operador A en $C_0(E)$ satisface el *principio del máximo positivo* si cada vez que $f \in \mathcal{D}(A)$, $x_0 \in E$ y $\sup f(x) = f(x_0) \geq 0$ se tiene que $Af(x_0) \leq 0$.

Lema 1.12 (Lema 4.2.1 en [11]) *Si S es localmente compacto, todo operador lineal sobre $C_0(S)$ que satisface el principio del máximo positivo es disipativo.*

Finalmente, se tiene el siguiente resultado.

Teorema 1.13 (Corolario 4.2.8 en [11]) *Sean S localmente compacto y A un operador lineal en $C_0(S)$ satisfaciendo el principio del máximo positivo. Para cada $x \in S$ existe un proceso de Feller X^x con distribución inicial δ_x y trayectorias cádlág si y solamente si A es conservativo.*

1.2.1. Problema de Martingala

Asociadas a un proceso de Feller X , existen diversas martingalas.

Proposición 1.14 (Proposición VII.1.6 en [19]) *Sea $f \in \mathcal{D}(A)$. Entonces, el proceso definido por*

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Af(X_s) ds$$

es una \mathbb{P}_x -martingala, para todo $x \in S$. Más aún, si $Af \equiv 0$ (resp. $Af \leq 0$, $Af \geq 0$) entonces $f(X_t)$ es una martingala (resp. submartingala, sobremartingala).

Resulta que esta proposición tiene una recíproca, la que caracteriza el generador de X .

Proposición 1.15 (Proposición VII.1.7 en [19]) *Sea $f \in C_0(S)$. Si existe $g \in C_0(S)$ tal que*

$$f(X_t) - f(X_0) - \int_0^t g(X_s) ds$$

es una \mathbb{P}_x -martingala para todo $x \in S$, entonces $f \in \mathcal{D}(A)$ y $Af = g$.

Por último, se puede extender esto al caso en que las funciones tienen una dependencia temporal.

Proposición 1.16 *Sea $f : \mathbb{R}_+ \times C_0(S) \rightarrow \mathbb{R}$ una función tal que $f(t, \cdot) \in \mathcal{D}(A)$ y $f \in B([0, t] \times S)$ para todo $t \geq 0$. Entonces*

$$f(t, X_t) - f(0, X_0) - \int_0^t Af(s, X_s) ds$$

1.3. COUPLING

es una \mathbb{P}_x -martingala para todo $x \in S$. Aquí, el operador \mathcal{A} definido por

$$\mathcal{A}f(t, x) = Af(t, \cdot)(x) + \frac{\partial f}{\partial t}(t, x)$$

es el generador espacio-temporal de X .

Observemos que todos estos resultados pueden interpretarse de la siguiente forma. Si $f \in \mathcal{D}(A)$, entonces

$$f(X_t) = f(X_0) + \int_0^t Af(X_s) ds + M_t^f$$

con M_t^f una martingala nula en 0. Es por esto que este tipo de resultados también son llamados *fórmulas de Itô*.

Una aplicación que será de interés más adelante es el resultado siguiente, válido para cualquier proceso de Feller X . Se anota

$$T_D = \inf\{t \geq 0 : X_t \in D\}$$

al tiempo de entrada al conjunto cerrado $D \in \mathcal{B}(S)$.

Lema 1.17 *Sea $D \in \mathcal{B}(S)$ un conjunto cerrado. Supongamos que existen $\alpha > 0$ y una función $V : D^c \rightarrow [1, \infty)$ tal que $V \in \mathcal{D}(A)$ y*

$$AV(x) \leq -\alpha V(x).$$

Entonces

$$\sup_{x \in S} \mathbb{E}_x(e^{\alpha T_D}) \leq \sup_{x \in D^c} V(x). \quad (1.10)$$

DEMOSTRACIÓN. Claramente, si $x \in D$ entonces $\mathbb{P}_x(T_D = 0) = 1$ por lo que la expresión de la izquierda en (1.10) es igual a

$$\sup_{x \in D^c} \mathbb{E}_x(e^{\alpha T_D}).$$

La función $f(t, x) = e^{\alpha t} V(x)$ es acotada en intervalos finitos, $f(t, \cdot) \in \mathcal{D}(A)$ para todo $t \geq 0$ y

$$\mathcal{A}f(t, x) = e^{\alpha t} AV(x) + \alpha e^{\alpha t} V(x) \leq 0.$$

Como conclusión, el proceso $Y_t = e^{\alpha t} V(X_t)$ es una submartingala positiva. Usando el Teorema Opcional de Doob y el hecho que $V \geq 1$ se obtiene

$$\mathbb{E}_x(e^{\alpha T_D}) \leq \mathbb{E}_x(Y_{T_D}) \leq \mathbb{E}_x(Y_0) = V(x)$$

para cualquier $x \in D^c$, lo que concluye la demostración. \square

1.3. Coupling

Sea (E, \mathcal{E}) un espacio medible y $\mu, \nu \in \mathcal{P}(E)$. Un coupling entre μ y ν es un par de variables aleatorias (X, Y) definidas en un espacio de probabilidad común $(\Omega, \mathcal{F}, \mathbb{P})$ con valores en $E \times E$, tal que X tiene distribución μ e Y tiene distribución ν , esto es,

$$\mathbb{P}(X \in A) = \mu(A), \quad \mathbb{P}(Y \in B) = \nu(B)$$

para todo $A, B \in \mathcal{E}$. Se denota por $\Pi(\mu, \nu)$ a la colección de todos los couplings de μ y ν . Es importante observar que siempre $\Pi(\mu, \nu)$ pues se puede tomar X e Y independientes con las distribuciones correspondientes, por lo que $\Pi(\mu, \nu)$ es no vacío.

La distancia en variación total entre μ y ν está definida por

$$\|\mu - \nu\|_{VT} = \frac{1}{2} \sup_{A \in \mathcal{E}} |\mu(A) - \nu(A)|.$$

Si μ y ν son absolutamente continuas respecto de una medida de referencia λ en \mathcal{E} (por ejemplo, $\lambda = \mu + \nu$) con densidades f y g respectivamente, entonces

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \int_E |f(x) - g(x)| \lambda(dx).$$

Lema 1.18 *Si la diagonal*

$$\Delta \doteq \{(x, x) : x \in E\}$$

es medible, esto es, $\Delta \in \mathcal{E} \otimes \mathcal{E}$, se tiene la relación dual

$$\|\mu - \nu\|_{VT} = \inf_{(X, Y) \in \Pi(\mu, \nu)} \mathbb{P}(X \neq Y).$$

La hipótesis del Lema anterior se satisface si, por ejemplo, (E, \mathcal{E}) es un espacio métrico Polaco.

1.3.1. Coupling Maximal

Dado un coupling $(X, Y) \in \Pi(\mu, \nu)$ definido en $(\Omega, \mathcal{F}, \mathbb{P})$, un conjunto $C \in \mathcal{F}$ se dice *evento de coupling para (X, Y)* si $C \subset \{X = Y\}$. Si las leyes de X e Y tiene densidad con respecto a una medida de referencia λ con densidades f y g respectivamente, es sencillo ver que se tiene la desigualdad

$$\mathbb{P}(C) \leq \int (f \wedge g)(x) \lambda(dx)$$

para cualquier evento de coupling C para (X, Y) .

Claramente, el evento $C = \{X = Y\}$ es un evento de coupling para (X, Y) . En particular, se tiene el siguiente resultado

Proposición 1.19 *Dadas dos medidas de probabilidad $\mu, \nu \in \mathcal{P}(E)$, existe un coupling $(X, Y) \in \Pi(\mu, \nu)$ tal que*

$$\|\mu - \nu\|_{TV} = \mathbb{P}(X \neq Y).$$

En particular, si X e Y tienen densidades f y g con respecto a una medida común λ , entonces

$$\mathbb{P}(X = Y) = \int_E (f \wedge g)(x) \lambda(dx). \quad (1.11)$$

1.3.2. Coupling entre procesos

Sea $(P_t)_{t \geq 0}$ una función de transición. Un coupling de un proceso de Markov con semigrupo P partiendo de x y un proceso de Markov con semigrupo P partiendo de y es un par de procesos (X^x, X^y) definidos en el mismo espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P}_{x,y})$, tal que X^x es un proceso de Markov con semigrupo P partiendo de x y X^y es un proceso de Markov con semigrupo P partiendo de y . De esta definición es fácil ver que (X^x, X^y) es un coupling si y solamente si el par (X_t^x, X_t^y) es un coupling de $P_t(x, \cdot)$ y $P_t(y, \cdot)$ en el sentido del párrafo anterior. Observemos que un coupling de dos procesos de Markov partiendo de x e y respectivamente no tiene por qué ser un proceso de Markov en sí mismo.

Si denotamos

$$T^* = \inf\{t \geq 0 : X_t^x \neq X_t^y\}$$

entonces en virtud del Lema precedente se tiene la desigualdad

$$\|P_t(x, \cdot) - P_t(y, \cdot)\|_{VT} \leq \mathbb{P}_{x,y}(X_t^x \neq X_t^y) = \mathbb{P}_{x,y}(T^* > t). \quad (1.12)$$

De esto es fácil deducir

Lema 1.20 *Supongamos que para cada $x, y \in E$ existe un coupling (X^x, X^y) . Entonces, para cualquier par de distribuciones iniciales $\mu, \nu \in \mathcal{P}(E)$ tales que*

$$\int_{E \times E} \mathbb{P}_{x,y}(T^* > t) \mu \otimes \nu(dx, dy) < \infty$$

se tiene que

$$\|\mu P_t - \nu P_t\|_{VT} \leq \int_{E \times E} \mathbb{P}_{x,y}(T^* > t) \mu \otimes \nu(dx, dy).$$

DEMOSTRACIÓN. Para $A \in \mathcal{E}$ se tiene, por definición

$$|\mu P_t(A) - \nu P_t(A)| = \left| \int_E P_t(x, A) \mu(dx) - \int_E P_t(y, A) \nu(dy) \right|.$$

Por definición de la medida producto,

$$\mu P_t(A) = \int_E P_t(x, A) \mu(dx) = \int_{E \times E} P_t(x, A) \mu \otimes \nu(dx, dy)$$

y similarmente con νP_t . Luego, por medio de la desigualdad triangular se obtiene que

$$|\mu P_t(A) - \nu P_t(A)| \leq \int_{E \times E} |P_t(x, A) - P_t(y, A)| \mu \otimes \nu(dx, dy).$$

Utilizando la definición de la distancia en variación total y la desigualdad (1.12) es posible deducir que el lado derecho de esta inecuación es a lo más

$$\int_{E \times E} \mathbb{P}_{x,y}(T^* > t) \mu \otimes \nu(dx, dy).$$

□

En particular, si

$$\int_{E \times E} \mathbb{E}_{x,y}(\varphi(T^*)) \mu \otimes \nu(dx, dy) < \infty$$

para alguna función creciente $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}$ tal que $\varphi(t) \rightarrow \infty$ cuando $t \rightarrow \infty$ se tendrá, utilizando la desigualdad de Markov, que

$$\int_{E \times E} \mathbb{P}_{x,y}(T^* > t) \mu \otimes \nu(dx, dy) \leq \varphi(t)^{-1} \int_{E \times E} \mathbb{E}_{x,y}(\varphi(T^*)) \mu \otimes \nu(dx, dy),$$

es decir, la distancia en variación total entre μP_t y νP_t decae como $\varphi(t)$ cuando $t \rightarrow \infty$. Más precisamente tendremos que existe una constante $C > 0$ tal que $\|\mu P_t - \nu P_t\|_{VT} \leq C\varphi(t)^{-1}$ para todo $t > t'$ con $\varphi(t') > 0$. Si $\pi \in \mathcal{P}(E)$ es una distribución estacionaria para P , esto dice que existe una tasa de convergencia al equilibrio del orden de φ .

Supongamos que (X^x, X^y) es un coupling Markoviano. Por lo discutido anteriormente, las leyes marginales de X_t^x e X_t^y son $P_t(x, \cdot)$ y $P_t(y, \cdot)$ respectivamente. Si denotamos por \tilde{P} al semigrupo asociado al par (X^x, X^y) , entonces esto se traduce en las identidades

$$\tilde{P}_t((x, y), A \times E) = P_t(x, A), \quad \tilde{P}_t((x, y), E \times B) = P_t(y, B)$$

para todo $A, B \in \mathcal{E}$. En particular, si f es una función que depende sólo de una variable (digamos $f = f(x)$) entonces

$$\tilde{P}_t f(x, y) = P_t f(x)$$

para todo $x, y \in E$. De esto sigue que si \tilde{A} es el generador de \tilde{P} y A el de P entonces

$$\tilde{A}f(x, y) = Af(x) \tag{1.13}$$

para toda función suave f que depende sólo de x .

Por último, se tiene que (X_t^x, X_t^y) es un coupling independiente de $P_t(x, \cdot)$ y $P_t(y, \cdot)$ si y solamente si

$$\tilde{A}f(x, y) = Af(\cdot, y)(x) + Af(x, \cdot)(y).$$

Esto se deduce fácilmente notando que en este caso $\tilde{P}_t((x, y), A \times B) = P_t(x, A)P_t(y, B)$ y aplicando \tilde{A} a f de la forma $f(x, y) = g(x)h(y)$.

1.4. Procesos de Markov deterministas por pedazos

Una clase particular de procesos de Markov son los llamados deterministas por pedazos, los que fueron introducidos por M.H.A Davis en 1984 (ver [7, 8, 15]) como una forma describir diversos modelos que no son de forma difusiva. La importancia de este tipo de procesos es que sus trayectorias no necesariamente son continuas a diferencia de las difusiones. Aún así, es posible construir un cálculo análogo al de Itô en este caso. Miembros representativos de esta clase son, por ejemplo, las vidas residual y acumulada asociadas a un proceso de renovación ([1, Capítulo V]), el llamado *Proceso TCP* ([6, 18]) y muchos otros modelos.

Los procesos de Markov deterministas por pedazos se construyen a partir de tres objetos. Se considera un flujo continuo $\Phi : \mathbb{R}_+ \times E \rightarrow \mathbb{R}_+$ en E , una función $\lambda : E \rightarrow \mathbb{R}_+$ llamada

1.4. PROCESOS DE MARKOV DETERMINISTAS POR PEDAZOS

intensidad y un núcleo de transición K en E . Se denotará $\Phi_t x$ al valor de Φ en (t, x) . Es necesario recordar que Φ es un flujo si satisface la condición de semigrupo $\Phi_{t+s}x = \Phi_t(\Phi_s x)$, para todo $t, s \geq 0$ y $x \in E$. La intensidad satisface

$$\int_0^\infty \lambda(\Phi_s x) ds = \infty \quad (1.14)$$

y para cada $x \in E$ existe $t(x) > 0$ tal que

$$\int_0^{t(x)} \lambda(\Phi_s x) ds < \infty. \quad (1.15)$$

Intuitivamente, la condición (1.14) asegura que el sistema no puede quedarse estancado en ningún estado y la condición (1.15) dice que se debe permanecer al menos una cantidad de tiempo en cada estado. Más formalmente, estas condiciones están relacionadas con la no explosión y la no extinción de X . Por comodidad se denota

$$\Lambda(t, x) = \int_0^t \lambda(\Phi_s x) ds \quad (1.16)$$

Notemos que la aplicación $t \mapsto \Lambda(t, x)$ es creciente, para cada $x \in E$ fijo.

Comenzando desde la posición $X_0 = x$, el proceso sigue el flujo determinista durante un intervalo de tiempo S_1 , escogido de forma que $\Lambda(S_1, x) = E_1$, en donde $E_1 \sim \mathcal{E}(1)$ es una v.a. independiente de X . Luego, salta a la posición y con distribución $K(\Phi_{S_1^-} x, \cdot)$, esto es,

$$\mathbb{P}_x(X_{S_1} \in A) = K(\Phi_{S_1^-} x, B), \quad B \in \mathcal{E}.$$

Este mecanismo es repetido a partir de la posición X_{S_1} . El tiempo

$$T_n^x = S_1 + \cdots + S_n \quad (1.17)$$

corresponde al tiempo real del n -ésimo salto.

De la definición, es claro que $\mathbb{P}_x(S_1 > t) = e^{-\Lambda(t, x)}$, y en general $\mathbb{P}_x(S_n > t \mid X_{T_n}) = e^{-\Lambda(t, X_{T_n})}$. Usando el Teorema de Convergencia Dominada y la condición (1.14) es fácil ver que $\mathbb{P}_x(S_n < \infty) = 1$ para todo $x \in E$ y $n \in \mathbb{N}$. De la misma forma, es posible observar que $\mathbb{P}_x(S_n > t(x)) > 0$.

Por último, de la construcción es directo deducir que la densidad conjunta de X_{T_1} y T_1 está dada por

$$\mathbb{P}_x(X_{T_1} \in dy, T_1 \in ds) = \lambda(\Phi_s x) e^{-\Lambda(s, x)} K(\Phi_s x, dy) ds \quad (1.18)$$

Definición 1.21 Una función $f : E \rightarrow \mathbb{R}$ se dice *trayectorialmente continua* (resp. *diferenciable*) si $t \mapsto f(\Phi_t x)$ es continua (resp. diferenciable) para todo $x \in E$. Se denota por $C_p(E)$ (resp. $D_p(E)$) a la clase de todas las funciones trayectorialmente continuas (resp. diferenciables).

De ahora en adelante se supondrá que $x \mapsto \lambda(x)Kf(x) \in C_p(E)$ para toda función f acotada y trayectorialmente continua. Si $f \in D_p(E)$, el límite

$$Gf(x) \doteq \lim_{h \downarrow 0} \frac{f(\Phi_h x) - f(x)}{h}$$

existe para todo $x \in E$. Más aún, en virtud de la propiedad de semigrupo de Φ se tiene que

$$\frac{d}{dt}f(\Phi_t x) = Gf(\Phi_t x).$$

Sea el operador $A : \mathcal{D}(A) \subseteq B(E) \rightarrow B(E)$ definido por

$$Af(x) = Gf(x) + \lambda(x) \int_E (f(y) - f(x)) K(x, dy), \quad (1.19)$$

actuando en el dominio $\mathcal{D}(A)$ consistente en las funciones $f \in D_p(E)$ acotadas tales que $Gf \in C_p(E)$ y $Af \in B(E)$.

A continuación se presenta una versión de la fórmula de Itô para procesos de Markov deterministas por pedazos. Se dirá que una función $f : \mathbb{R}_+ \times E \rightarrow \mathbb{R}$ es acotada en intervalos finitos si $f \in B([0, t] \times E)$ para todo $t \geq 0$.

Proposición 1.22 (Proposición 7.7.1 en [15]) *Sea $f : \mathbb{R}_+ \times E$ acotada en intervalos finitos tal que $t \mapsto f(t, x)$ es de clase $C^1(\mathbb{R}_+)$ para todo $x \in E$, y $x \mapsto f(t, x)$ pertenece a $D_p(E)$ para todo $t \geq 0$. Entonces*

$$f(t, X_t) = f(0, X_0) + \int_0^t \mathcal{A}f(s, X_s) ds + M_t^f, \quad (1.20)$$

donde $\mathcal{A}f(t, x) = \frac{\partial f}{\partial t}(t, x) + Af(t, \cdot)(x)$ es el generador espacio-temporal de X y M_t^f es una martingala nula en 0 y A está definido en (1.19).

Asociada a un proceso de Markov determinista por pedazos X existe una cadena de Markov \hat{X} con valores en \mathbb{R}_+ , llamada *cadena incrustada*, la que está definida por

$$\hat{X}_n = X_{T_n}. \quad (1.21)$$

De la ecuación (1.18) es fácil ver que el núcleo de transición asociado a \hat{X} está dado por

$$Rf(x) = \int_0^\infty \lambda(\Phi_s x) Kf(\Phi_s x) e^{-\Lambda(s, x)} ds = \mathbb{E}_x(Kf(\Phi_{T_1} x)).$$

1.5. La clase Incremento Aditivo Decremento Multiplicativo

El nombre de esta clase proviene de las siglas en inglés para *Additive-Increase Multiplicative-Decrease*. Como su nombre lo indica, esta subclase de los procesos de Markov deterministas por pedazos está caracterizada por el hecho de que su incremento determinista es aditivo y sus saltos son de tipo multiplicativo. Más precisamente, X pertenece a la clase AIMD si toma valores en \mathbb{R}_+ , su flujo Φ es aditivo y creciente, esto es, la función $t \mapsto \Phi_t x$ es creciente y $\Phi_{t+s} x = \Phi_t x + \Phi_s x$ para todo $x \in \mathbb{R}_+$ y $t, s \geq 0$; y su núcleo de transición tiene la forma

$$K(x, A) = H(x^{-1}A), \quad A \in \mathcal{B}(\mathbb{R}_+) \quad (1.22)$$

1.5. LA CLASE INCREMENTO ADITIVO DECREMENTO MULTIPLICATIVO

para alguna medida H concentrada en el intervalo $[0, 1)$. Se denota por $\mu = \int_0^1 h H(dh)$ a la media de H . Naturalmente $\mu \in [0, 1)$. Más aún, a veces se utilizará la letra H para denotar a la función acumulada de la medida H , esto es, $H : [0, 1] \rightarrow [0, 1]$ es la función creciente y continua por la derecha definida por $H(h) = H([0, h))$.

La propiedad de aditividad de Φ junto con la continuidad de $t \mapsto \Phi_t x$ para todo $x \in \mathbb{R}_+$ implican que existe una constante $\beta > 0$ tal que

$$\Phi_t x = x + \beta t$$

para todo $(t, x) \in \mathbb{R}_+^2$. Definiendo \tilde{X} por $\tilde{X}_t = X_{t/\beta}$ es fácil ver que el flujo $\tilde{\Phi}$ correspondiente a \tilde{X} es de la forma $\tilde{\Phi}_t x = x + t$. Es por esto que se supondrá de ahora en adelante, sin pérdida de generalidad, que $\beta = 1$. Usando la forma de Φ , se tiene que la función Λ definida en (1.16) tiene la forma $\Lambda(t, x) = \int_0^t \lambda(x + s) ds$ y, haciendo el cambio de variable lineal $u = x + s$ se deduce la desigualdad

$$\Lambda(t, x) \leq \Lambda(t + x, 0). \quad (1.23)$$

Por otra parte, dado que H se concentra en $[0, 1)$, de la ecuación (1.22) se puede ver que el núcleo $K(x, \cdot)$ se concentra en el intervalo $[0, x)$. Más específicamente, para cualquier función $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ se tiene que

$$Kf(x) = \int_0^1 f(hx) H(dh) = \mathbb{E}(f(Qx)),$$

en donde Q es una v.a. con función de distribución acumulada H .

Dada la forma del flujo Φ , una función será trayectorialmente continua (resp. diferenciable) si y solamente si es continua (resp. diferenciable) en el sentido clásico. Con todo esto, el generador infinitesimal de X toma la forma

$$Af(x) = f'(x) + \lambda(x) \int_0^1 (f(hx) - f(x)) H(dh) \quad (1.24)$$

definido en el dominio $\mathcal{D}(A) \subset C^1(\mathbb{R}_+)$ consistente en las funciones continuamente diferenciables tales que $Af \in B(\mathbb{R}_+)$.

1.5.1. Algunos ejemplos

A continuación se muestra que ciertos procesos pertenecen a la clase AIMD, y se calcula explícitamente su tasa y núcleo de saltos.

Procesos de Renovación

Un proceso de renovación S es una sucesión de variables aleatorias $(S_n)_{n \geq 0}$ tal que $(S_n)_{n \geq 1}$ es i.i.d. con distribución común F , y es independiente de S_0 , quien tiene distribución F_0 . Por comodidad se define $T_n = S_0 + S_1 + \cdots + S_n$. Si $S_0 \equiv 0$ c.s. se dirá que S es puro, y en caso contrario se dirá que está retardado, con distribución de retraso F_0 .

Asociado a S , hay un proceso de conteo $(N_t)_{t \geq 0}$ definido por

$$N_t = \inf\{n \geq 0 : T_n \geq t\}.$$

Este proceso es creciente y continuo a la derecha. Intuitivamente, N_t es la cantidad de *eventos* ocurridos hasta el tiempo $t \geq 0$. A partir de N_t podemos definir la *edad actual* de S mediante $X_t = t - T_{N_t-1}$, vale decir, el tiempo transcurrido desde el último evento. Para una revisión de las definiciones y resultados fundamentales ver el Capítulo V de [1].

Proposición 1.23 (Proposición V.1.5 en [1]) *El proceso $(X_t)_{t \geq 0}$ es de Feller.*

Veamos que $(X_t)_{t \geq 0}$ pertenece a la clase AIMD. De la definición, es claro que el flujo determinista es de la forma $\Phi_t x = x + t$. Además, usando la propiedad de Markov se tiene, para $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ de clase C^1

$$\begin{aligned} \mathbb{E}_x(f(X_t)) &= \mathbb{E}_x(f(X_t)\mathbf{1}_{S_0 > t}) + \mathbb{E}_x(f(X_t)\mathbf{1}_{S_0 \leq t}) \\ &= \mathbb{E}_x(f(x+t)\mathbf{1}_{\{S_0 > t\}}) + \mathbb{E}_x(\mathbb{E}_x(f(X_t) \mid \mathcal{F}_{S_0})\mathbf{1}_{\{S_0 \leq t\}}) \\ &= \mathbb{E}_x(f(x+t)\mathbf{1}_{\{S_0 > t\}}) + \mathbb{E}_x(\mathbb{E}_0(f(X_{t-u}) \mid_{u=S_0})\mathbf{1}_{\{S_0 \leq t\}}) \end{aligned}$$

Pero como $\mathbb{P}_x(S_0 > t) = \frac{\bar{F}(x+t)}{\bar{F}(x)}$, el término de la derecha es igual a

$$\frac{1}{\bar{F}(x)} \int_x^{x+t} \mathbb{E}_0(f(X_{t-u})) F(du)$$

y con esto

$$\mathbb{E}_x(f(A_t)) = \frac{1}{\bar{F}(x)} \left[f(x+t) + \int_x^{x+t} \mathbb{E}_0(f(X_{t-u})) F(du) \right].$$

Luego, si F es diferenciable se tendrá que

$$\begin{aligned} Af(x) &= \lim_{t \downarrow 0} \frac{\mathbb{E}_x(f(X_t)) - f(x)}{t} \\ &= f'(x) + \frac{F'(x)}{1 - F(x)} (f(0) - f(x)) \end{aligned}$$

para todo $f \in C^1(\mathbb{R}_+)$.

Proposición 1.24 *La edad actual asociada a un proceso de renovación con distribución de espera F diferenciable pertenece a la clase AIMD, con tasa $\lambda(x) = \frac{F'(x)}{1-F(x)}$ y núcleo de transición $H(dh) = \delta_0(dh)$.*

En la Ilustración 1.1 se presenta una trayectoria de un proceso de renovación con $F(x) = 1 - e^{-x}$ y distribución inicial $\mathcal{L}(X_0) = \mathcal{E}(1)$.

Es sabido que el proceso X es ergódico, y su distribución invariante tiene densidad

$$F_0(x) = \frac{1}{\mu} (1 - F(x)),$$

en donde $\mu = \int_0^\infty x F(dx)$ es el tiempo medio de espera.

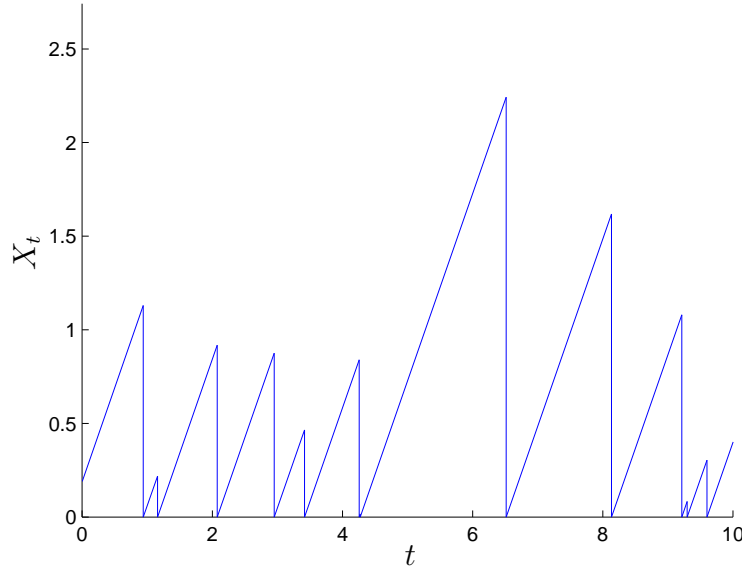


Ilustración 1.1: Trayectoria de la edad actual asociada a un proceso de renovación, con $F(x) = 1 - e^{-x}$ y $\mathcal{L}(X_0) = \mathcal{E}(1)$.

Un resultado clásico respecto de la ergodicidad exponencial de este tipo de procesos es el siguiente. Se dice que F tiene una *componente G* si $0 \neq G \leq F$. Si F^{*n} tiene una componente absolutamente continua para algún n , diremos que F es *spread-out*. Aquí, F^{*n} es la convolución de F con síg misma, n veces; esto es,

$$F^{*0} = F, \quad F^{*(n+1)} = F \star F^{*n}.$$

Teorema 1.25 (Teoremas VII.2.7 y VII.2.10 en [1]) *El proceso X admite coupling si y solamente si $\mu < \infty$, y F es no aritmética y spread-out. Más aún, si $\int_0^\infty e^{\eta x} F(dx) < \infty$ para algún $\eta > 0$, se tiene que*

$$\|\mathbb{P}_0(X_t \in \cdot) - F_0\| = O(e^{-\varepsilon t})$$

para algún $\varepsilon > 0$.

El proceso TCP

El proceso TCP modela uno de los protocolos de transmisión más utilizados en el Internet. El protocolo TCP regula el tráfico en una conexión dada. Para esta conexión, la cantidad máxima de paquetes de datos que se puede enviar a través del canal está dada por una v.a. discreta W , llamada *ventana de congestión*. Si se envían satisfactoriamente todos los paquetes de datos, W es incrementada en uno. En caso contrario, es multiplicada por un número $\delta \in [0, 1)$. Bajo el supuesto de que las pérdidas son aleatorias i.i.d. exponenciales, un correcto escalamiento de este algoritmo da un proceso de Markov a tiempo continuo X con generador

$$Af(x) = f'(x) + x \int_0^1 (f(hx) - f(x)) H(dh),$$

en donde H es una medida concentrada en el intervalo $[0, 1)$, como se muestra en [10, 13, 18]. En [13] se prueba además que posee una única medida invariante π concentrada en $(0, \infty)$

con densidad

$$x \mapsto \frac{\sqrt{2/\pi}}{\prod_{n \geq 0} (1 - 2^{-(2n+1)})} \sum_{n \geq 0} \frac{(-1)^n 2^{2n}}{\prod_{k=1}^n (2^{2k} - 1)} e^{-2^{2n-1} x^2}.$$

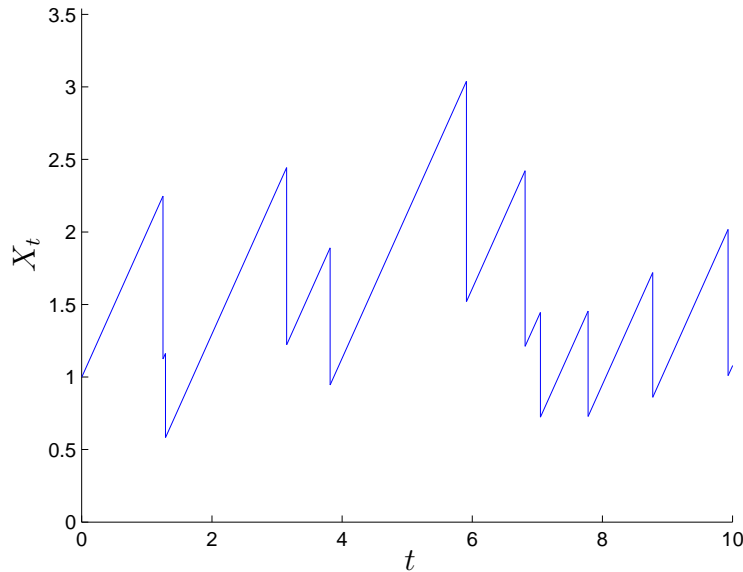


Ilustración 1.2: Trayectoria del proceso TCP con $X_0 \equiv 1$

Se tiene así que este proceso pertenece a la clase AIMD con $\lambda(x)$ y H es general.

Numerosos resultados de ergodicidad exponencial han sido probados en este caso emblemático. En 2009, Chafaï *et al.* probaron en [6] que se tiene una tasa de convergencia polinomial en la distancia de Wasserstein W_1 , en el caso en que el factor H es constante.

Teorema 1.26 (Teorema 2.5 en [6]) *Supongamos que $H = \delta_h$, con $h \in (0, 1)$. Entonces para cualquier $t \geq 0$ y leyes iniciales μ, ν con primer momento finito, existe un coupling (X, Y) tal que*

$$W_1(\mu P_t, \nu P_t) \leq \frac{W_1(\mu, \nu)}{1 + (1 + h)W_1(\mu, \nu)t}.$$

En este paper se deja abierta la pregunta de si es posible encontrar una tasa exponencial, y otras posibles distancias; por ejemplo, W_p para $p \geq 1$ o variación total. Esta pregunta fue respondida de forma afirmativa por Bardet *et al.* en 2011. Concretamente, probaron

Teorema 1.27 (Teorema 1.4 en [4]) *Existe una constante $\hat{\gamma}$ tal que para cualquier $\gamma < \hat{\gamma}$, $p \geq 1$ y $t_0 > 0$, existe una constante $C(p, \gamma, t_0)$ tal que, para cualquier par de leyes iniciales μ, ν y $t \geq 0$,*

$$W_p(\nu P_t, \mu P_t) \leq C \exp\left(-\frac{\gamma}{p}t\right).$$

Teorema 1.28 (Teorema 1.5 en [4]) *Para cualquier $\gamma < \hat{\gamma}$ y $t_0 > 0$ existe una constante*

1.5. LA CLASE INCREMENTO ADITIVO DECREMENTO MULTIPLICATIVO

$C > 0$, tal que para cualquier par de leyes iniciales μ y ν y $t \geq t_0$,

$$\|\mu P_t - \nu P_t\|_{VT} \leq C \exp\left(-\frac{2\gamma}{3}t\right),$$

en donde $\hat{\gamma}$ es el mismo del Teorema anterior.

Cabe destacar que en estos resultados, la constante $\hat{\gamma} \sim 0,12$ es explícita.

Capítulo 2

Resultados

Este capítulo está dedicado a la presentación de los resultados obtenidos en este trabajo de memoria. En primera instancia se desarrollan resultados generales relacionados con la convergencia al equilibrio de Procesos de Markov deterministas por pedazos pertenecientes a la subclase AIMD. Posteriormente, se aplican estos resultados a diversos ejemplos provenientes de distintas áreas y que ya han sido estudiados mediante otras técnicas. Finalmente se presentan algunas simulaciones numéricas a modo de comparación entre lo obtenido aquí y los resultados ya existentes en la literatura.

En el Teorema principal se describe la velocidad de convergencia de estos procesos; específicamente, se prueba que bajo ciertas condiciones hay una convergencia exponencial al equilibrio, encontrándose tasas explícitas. Para la demostración de este hecho, es útil estudiar los tiempos de entrada a intervalos de la forma $[0, x_0]$ y $[x_0, \infty)$. En particular, se analizan dichos tiempos de entrada en el caso en que la intensidad del proceso es una constante positiva y mediante una técnica de dominación a través de couplings se obtienen estimaciones en el caso general. Posteriormente se prueban cotas para la probabilidad de coalescencia de un coupling particular, y finalmente se construye uno con las propiedades deseadas.

Dado $u \geq 0$, denotaremos por X^u a la versión de X , definida en el espacio $(\Omega, \mathcal{F}, \mathbb{P})$ tal que $\mathbb{P}(X_0^u = u) = 1$

2.1. Tiempos de retorno

2.1.1. El caso de tasa constante

En primer lugar se estudia el caso en que el proceso de interés tiene intensidad constante $\lambda > 0$, la que estará fija de ahora en adelante. Consideramos entonces un proceso Z a valores en \mathbb{R}_+ cuyo generador infinitesimal está dado por

$$Af(u) = f'(u) + \lambda \int_0^1 (f(hu) - f(u)) H(dh) \quad (2.1)$$

para $f \in C_0^1(\mathbb{R}_+)$ y $u \geq 0$.

2.1. TIEMPOS DE RETORNO

Dados $x_0 \geq 0$ y $s \geq 0$ fijos, definimos

$$\tau'(s) = \inf\{t \geq 0 : Z_{t+s} \leq x_0\} \quad (2.2)$$

$$\sigma'(s) = \inf\{t \geq 0 : Z_{t+s} \geq x_0\} \quad (2.3)$$

El tiempo $\tau'(s)$ (resp. $\sigma'(s)$) corresponde al tiempo relativo, medido desde s , en el que Z pasa bajo (resp. sobre) el nivel x_0 .

La igualdad

$$\{\tau'(s) > r\} = \bigcap_{\substack{u \leq r \\ u \in \mathbb{Q}_+}} \{Z_{s+u} > x_0\}$$

muestra que para cada $s \in \mathbb{R}_+$, $\tau'(s)$ es un tiempo de parada con respecto a la filtración desplazada $\mathcal{F}_r^s = \mathcal{F}_{r+s}$. De forma análoga se puede ver que $\sigma'(s)$ es también tiempo de parada con respecto a \mathcal{F}^s . Más aún, de la propiedad de Markov fuerte (ecuación (1.9)) de Z se deduce que, para tiempos de parada $T \geq S$ finitos c.s.

$$\mathbb{E}[f(\tau(T)) \mid \mathcal{F}_S] = \mathbb{E}_{Z_S}[f(\tau(T - S))] \quad (2.4)$$

para cualquier función $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ medible y positiva.

En lo que sigue se denotará $\tau' = \tau'(0)$ y $\sigma' = \sigma'(0)$.

Lema 2.1 *Supongamos que $x_0 > 2[\lambda(1 - \mu)]^{-1}$. Entonces para todo $z > x_0$ y todo $0 < \gamma < x_0^{-1}(\lambda(1 - \mu)x_0 - 2)$ se tiene*

$$\mathbb{E}_z(e^{\gamma\tau'}) \leq 1 + \gamma z$$

DEMOSTRACIÓN. Consideremos la función $V(z) = 1 + \gamma z$. Directamente se obtiene

$$AV(z) = \gamma - \lambda\gamma(1 - \mu)z = \gamma(1 - \lambda(1 - \mu)z).$$

Buscamos $\alpha_\gamma > 0$ tal que

$$\gamma(1 - \lambda(1 - \mu)z) \leq -\alpha_\gamma(1 + \gamma z),$$

para todo $z > x_0$. Es fácil ver que esta se tiene esta desigualdad si y solamente si

$$\alpha_\gamma \leq \frac{\gamma(\lambda(1 - \mu)z - 1)}{1 + \gamma z} \doteq h(z), \quad \forall z > x_0.$$

Derivando el miembro derecho de esta expresión se tiene que, para cualquier $z > 0$,

$$\begin{aligned} h'(z) &= \frac{\lambda\gamma(1 - \mu)(1 + \gamma z) - \gamma^2(\lambda(1 - \mu)z - 1)}{(1 + \gamma z)^2} \\ &= \frac{\gamma^2 + \lambda\gamma(1 - \mu)}{(1 + \gamma z)^2} > 0, \end{aligned}$$

por lo que h es una función creciente. Luego, basta escoger

$$\alpha_\gamma = \inf_{z > x_0} h(z) = \frac{\gamma(\lambda(1 - \mu)x_0 - 1)}{1 + \gamma x_0} > 0.$$

Por otra parte, usando la hipótesis sobre γ se ve que $\alpha_\gamma \geq \gamma$.

Con esto, V satisface las hipótesis del Lema 1.17 por lo que

$$\mathbb{E}_z(e^{\gamma\tau'}) \leq \mathbb{E}_z(e^{\alpha_\gamma\tau'}) \leq 1 + \gamma z$$

para todo $z > x_0$. □

Observación En particular, usando la propiedad de Markov fuerte (1.9) se deduce que, para todo tiempo de parada T finito c.s. tal que $Z_T > x_0$, vale la siguiente estimación:

$$\mathbb{E}_z \left(e^{\gamma \tau'(T)} \right) \leq 1 + \gamma \mathbb{E}_z(Z_T).$$

Antes de proceder con el siguiente resultado, necesitamos el siguiente Lema, tomado de [3].

Lema 2.2 Sea (\mathcal{G}_n) una filtración, y sean (A_n) y (W_n) secuencias adaptadas de eventos y variables aleatorias no negativas. Sea $\theta \in (0, 1)$ y $\psi : I \rightarrow \mathbb{R}$ una función definida en el intervalo $I \ni 0$ tal que $\psi(0) = 0$ y $\psi(\lambda) < -\ln(1 - \theta)$ para todo $\lambda \in I$. Se define

$$N \doteq \inf\{n \geq 1 : \mathbf{1}_{A_n} = 1\}$$

y supongamos que para todo $n \geq 1$ y $\lambda \in I$ se tiene que

1. $\mathbb{E}(e^{\lambda W_n} \mid \mathcal{G}_{n-1}) \leq e^{\psi(\lambda)}$ y
2. $\mathbb{E}(e^{\lambda W_n} \mathbf{1}_{A_n^c} \mid \mathcal{G}_{n-1}) \leq (1 - \theta)e^{\psi(\lambda)}$.

Entonces, para todo $\lambda \in I$,

$$\mathbb{E} \left(\exp \left(\lambda \sum_{j=1}^N W_j \right) \right) \leq \mathbb{E} (e^{\psi(\lambda)G}) = \frac{\theta e^{\psi(\lambda)}}{1 - e^{\psi(\lambda)}(1 - \theta)},$$

donde G es una variable aleatoria geométrica de parámetro θ .

DEMOSTRACIÓN. Dado $K \in \mathbb{N}$, reemplazando (A_n) por (\tilde{A}_n) definido como $\tilde{A}_n = A_n$ si $n \leq K$ y $\tilde{A}_n = \Omega$ si $n \geq K + 1$, podemos suponer que $N = N \wedge K \leq K$ y usar el teorema de convergencia monótona para pasar al caso general. Usando 1. en la primera desigualdad y 2. en la segunda tenemos

$$\begin{aligned} \mathbb{E}(e^{\lambda \sum_{j=1}^N W_j} - 1) &= \sum_{n=1}^{\infty} \mathbb{E} \left(\left[e^{\lambda \sum_{j=1}^n W_j} - e^{\lambda \sum_{j=1}^{n-1} W_j} \right] \mathbf{1}_{\{N \geq n\}} \right) \\ &= \sum_{n=1}^{\infty} \mathbb{E} \left(e^{\lambda \sum_{j=1}^{n-1} W_j} \mathbb{E} (e^{\lambda W_n} - 1 \mid \mathcal{G}_{n-1}) \mathbf{1}_{\{N \geq n\}} \right) \\ &\leq \sum_{n=1}^{\infty} [e^{\psi(\lambda)} - 1] \mathbb{E} \left(e^{\lambda \sum_{j=1}^{n-1} W_j} \mathbf{1}_{A_{n-1}^c} \cdots \mathbf{1}_{A_1^c} \right) \\ &= \sum_{n=1}^{\infty} [e^{\psi(\lambda)} - 1] \mathbb{E} \left(\mathbb{E} \left(e^{\lambda W_{n-1}} \mathbf{1}_{A_{n-1}^c} \mid \mathcal{G}_{n-2} \right) e^{\lambda \sum_{j=1}^{n-2} W_j} \mathbf{1}_{A_{n-2}^c} \cdots \mathbf{1}_{A_1^c} \right) \\ &\leq \sum_{n=1}^{\infty} [e^{2\psi(\lambda)} - e^{\psi(\lambda)}] (1 - \theta) \mathbb{E} \left(e^{\lambda \sum_{j=1}^{n-2} W_j} \mathbf{1}_{A_{n-2}^c} \cdots \mathbf{1}_{A_1^c} \right) \end{aligned}$$

El hecho de que N es acotado justifica el intercambio de sumas con diferencias. Condicionando

2.1. TIEMPOS DE RETORNO

en \mathcal{G}_{n-3} e iterando el argumento da que

$$\begin{aligned} \mathbb{E}(e^{\lambda \sum_{j=1}^N W_j} - 1) &\leq \dots \leq \sum_{n=1}^{\infty} [e^{n\psi(\lambda)} - e^{(n-1)\psi(\lambda)}] (1 - \theta)^{n-1} \\ &= \sum_{n=1}^{\infty} [e^{n\psi(\lambda)} - e^{(n-1)\psi(\lambda)}] \mathbb{P}(G \geq n) \\ &= \mathbb{E}(e^{\psi(\lambda)G}). \end{aligned}$$

□

El caso $\lambda = 0$ en 2. corresponde al Lema A.6 en [1], por lo que este es un caso particular del Lema anterior. En lo que sigue, dado $x_0 > 0$ y G una variable aleatoria geométrica de parámetro $\theta = e^{-\lambda x_0}$, se denota

$$G_{x_0}^\lambda(\gamma) \doteq \mathbb{E}(e^{\gamma x_0 G}) = \frac{e^{(\gamma-\lambda)x_0}}{1 - e^{\gamma x_0}(1 - e^{-\lambda x_0})}. \quad (2.5)$$

Lema 2.3 Para todo $z < x_0$ y $0 < \gamma < -x_0^{-1} \ln(1 - e^{-\lambda x_0})$ se tiene

$$\mathbb{E}_z(e^{\gamma \sigma'}) \leq G_{x_0}^\lambda(\gamma)$$

DEMOSTRACIÓN. Sea $N = \inf\{n \in \mathbb{N} : \hat{Z}_n + S_{n+1} \geq x_0\}$, donde \hat{Z} es la cadena de Markov asociada a X , definida en (1.21). Claramente se tiene que $\mathbb{P}_z(\sigma \leq T_N^z \leq x_0 N) = 1$, esto pues por definición $Z_{(T_N^z)^-} \geq x_0$ y el hecho de que Z crece con pendiente igual a 1, donde T_N^z es el tiempo del N -ésimo salto de Z , definido en (1.17). Además, usando la propiedad de Markov de la cadena \hat{Z} se tiene que

$$\begin{aligned} \mathbb{P}_z(\hat{Z}_n + S_{n+1} \geq x_0 \mid \mathcal{F}_n) &= \mathbb{P}_{\hat{Z}_n}(Z_0 + S_1 \geq x_0) \\ &= e^{-\lambda(x_0 - \hat{Z}_n)} \\ &\geq e^{-\lambda x_0} \end{aligned}$$

de donde se deduce, en virtud del Lema 2.2, que

$$\mathbb{E}_z(e^{\gamma \sigma'}) \leq \mathbb{E}_z(e^{\gamma x_0 N}) \leq G_{x_0}^\lambda(\gamma)$$

□

Observación Dado que el lado derecho de la estimación anterior no depende de z , se deduce fácilmente, usando la propiedad de Markov fuerte, que la misma desigualdad es válida si reemplazamos σ por $\sigma'(T)$ para cualquier tiempo de parada T finito c.s. tal que $Z_T < x_0$, esto es,

$$\mathbb{E}_z(e^{\gamma \sigma'(T)}) \leq G_{x_0}^\lambda(\gamma).$$

2.1.2. Caso general

En esta sección se extienden los resultados obtenidos para el caso con intensidad constante al caso en que la intensidad λ satisface ciertas condiciones generales. Esto se realiza a través una técnica de *dominación* mediante couplings adecuados.

A partir de ahora, consideramos el proceso de Markov (X^x, X^y) con generador de la forma

$$\tilde{A}f(x, y) = \nabla \cdot f(x, y) + \beta(x, y) \int_{\mathbb{R}^2} (f(x', y') - f(x, y)) K((x, y), dx' dy'), \quad (2.6)$$

en donde el núcleo de transición K en \mathbb{R}_+^2 está dado por

$$Kf(x, y) = p_{x,y} \mathbb{E}(f(Qx, y)) + p_{y,x} \mathbb{E}(f(x, Qy)) + q_{x,y} \mathbb{E}(f(Qx, Qy))$$

para $f \geq 0$ medible y Q una v.a. con distribución H . Aquí, $p_{x,y}$ representa la probabilidad de que, al momento de que el coupling decida saltar en un tiempo $t > 0$ y $(X_{t-}^x, X_{t-}^y) = (x, y)$, la componente que se encuentra en x lo haga y la otra no. En símbolos,

$$p_{x,y} = K((x, y), [0, x] \times \{y\}).$$

Las cantidades $p_{y,x} = K((x, y), \{x\} \times [0, y])$ y $q_{x,y} = K((x, y), \Delta_{x,y}) = 1 - p_{x,y} - p_{y,x} = q_{y,x}$ pueden interpretarse de igual forma. Aquí $\Delta_{x,y} = \{(tx, ty) : t \in [0, 1]\}$.

Para que (X^x, X^y) sea un coupling de $P_t(x, \cdot)$ y $P_t(y, \cdot)$, la función de intensidad β que aparece en la ecuación (2.6) debe cumplir ciertas condiciones. Imponiendo las igualdades en (1.13) se deduce que

$$\lambda(x) = (1 - p_{y,x})\beta(x, y), \quad \lambda(y) = (1 - p_{x,y})\beta(x, y).$$

De aquí sigue que β debe ser simétrica, y satisfacer

$$\lambda(x) \vee \lambda(y) \leq \beta(x, y) \leq \lambda(x) + \lambda(y). \quad (2.7)$$

Es importante observar que en el caso en que $\beta(x, y) = \lambda(x) + \lambda(y)$ se tiene que

$$p_{x,y} = \frac{\lambda(x)}{\lambda(x) + \lambda(y)}, \quad p_{y,x} = \frac{\lambda(y)}{\lambda(x) + \lambda(y)}, \quad q_{x,y} = 0$$

y por lo tanto, reemplazando en (2.6), se ve que

$$\tilde{A}f(x, y) = Af(\cdot, y)(x) + Af(x, \cdot)(y),$$

esto es, (X^x, X^y) es un coupling independiente.

En el otro extremo, la elección $\beta(x, y) = \lambda(x) \vee \lambda(y)$ es la que maximiza la probabilidad de salto conjunto

$$q_{x,y} = \frac{\lambda(x) + \lambda(y)}{\lambda(x) \vee \lambda(y)} - 1 = \frac{\lambda(x) \wedge \lambda(y)}{\lambda(x) \vee \lambda(y)}.$$

2.1. TIEMPOS DE RETORNO

Hipótesis 1 De ahora en adelante se supondrá que existe $\delta > 0$ tal que

$$\eta \doteq \inf_{x > \delta} \lambda(x) > \frac{1}{1 - \mu}.$$

Esto es equivalente a suponer que

$$\liminf_{x \rightarrow \infty} \lambda(x) > \frac{1}{1 - \mu}.$$

Definimos $\bar{\eta} \doteq [\eta(1 - \mu)]^{-1} \in (0, 1)$, y fijamos $\varepsilon > 2\bar{\eta} \vee \delta$ y $\hat{\varepsilon} > \varepsilon + \delta$.

Como en la sección anterior consideremos, para $s \geq 0$,

$$\tau(s) = \inf\{t \geq 0 : X_{t+s}^x \vee X_{t+s}^y \leq \varepsilon\}, \quad (2.8)$$

$$\sigma(s) = \inf\{t \geq 0 : X_{t+s}^x \vee X_{t+s}^y \geq \hat{\varepsilon}\}, \quad (2.9)$$

$$\bar{\sigma}(s) = \inf\{t \geq 0 : X_{t+s}^x \wedge X_{t+s}^y \geq \delta\}. \quad (2.10)$$

Lema 2.4 Sean $x, y \in \mathbb{R}_+$ tales que $x \vee y > \varepsilon$. Entonces, para todo $0 < \gamma < \varepsilon^{-1}(\eta(1 - \mu)\varepsilon - 2)$ se tiene

$$\mathbb{E}_{x,y}(e^{\gamma\tau}) \leq 1 + \gamma(x \vee y). \quad (2.11)$$

DEMOSTRACIÓN. Consideremos Z^z un proceso partiendo desde $z \geq 0$, con generador (2.1) para $\lambda = \eta$, y el siguiente coupling entre (X^x, X^y) con generador (2.6) y Z^z . El proceso $((X^x, X^y), Z^z)$ sigue la dinámica dada por el generador

$$\begin{aligned} Bf(u, v, w) &= \nabla \cdot f(u, v, w) + (\beta(u, v) - \eta) \int_{\mathbb{R}_+^2} (f(u', v', w) - f(u, v, w)) K((u, v), du' dv') \\ &\quad + \eta \int_0^1 (f(hu, hv, hw) - f(u, v, w)) H(dh). \end{aligned}$$

si $u \vee v > \varepsilon$, e independientes en caso contrario, esto es,

$$\begin{aligned} Bf(u, v, w) &= \nabla \cdot f(u, v, w) + \beta(u, v) \int_{\mathbb{R}_+^2} (f(u', v', w) - f(u, v, w)) K((u, v), du' dv') \\ &\quad + \eta \int_0^1 (f(u, v, hw) - f(u, v, w)) H(dh) \end{aligned}$$

si $u \vee v \leq \varepsilon$. Con esto, si $z \geq x \vee y > \varepsilon$ entonces $Z_t^z \geq X_t^x \vee X_t^y$ en el intervalo $[0, \tau]$. Luego, si consideramos

$$\tau' = \inf\{t \geq 0 : Z_t^z \leq \varepsilon\}$$

claramente $\mathbb{P}_{x,y,z}(\tau \leq \tau') = 1$ siempre que $z \geq x \vee y$. Por lo tanto, utilizando el Lema 2.1

$$\begin{aligned} \mathbb{E}_{x,y}(e^{\gamma\tau}) &= \mathbb{E}_{x,y,x \vee y}(e^{\gamma\tau}) \\ &\leq \mathbb{E}_{x,y,x \vee y}(e^{\gamma\tau'}) \\ &= \mathbb{E}_{x \vee y}(e^{\gamma\tau'}) \\ &\leq 1 + \gamma(x \vee y). \end{aligned}$$

□

Hipótesis 2 A partir de este punto supondremos además que

$$M' \doteq \sup_{x \in [0, \hat{\varepsilon}]} \lambda(x) < \infty$$

y denotamos $M = 2M'$.

Lema 2.5 Sean $x, y \in \mathbb{R}_+$ tales que $x \wedge y < \delta$. Para todo $0 < \gamma < -\frac{1}{\delta} \ln(1 - e^{-M\delta})$ se tiene que

$$\mathbb{E}_{x,y}(e^{\gamma\bar{\sigma}}) \leq G_{\delta}^M(\gamma) \quad (2.12)$$

DEMOSTRACIÓN. Al igual que en la demostración del Lema 2.4, consideremos Z^z un proceso partiendo desde $z \geq y$ y con generador (2.1) para $\lambda = M$. Definiremos un coupling adecuado entre (X^x, X^y) y Z^z . El proceso $((X^x, X^y), Z^z)$ seguirá la dinámica dada por el generador

$$\begin{aligned} Bf(u, v, w) = & \nabla \cdot f(u, v, w) + \beta(u, v) \int (f(u', v', w') - f(u, v, w)) K'((u, v, w), du' dv' dw') \\ & + (M - \beta(u, v)) \int_0^1 (f(u, v, hw) - f(u, v, w)) H(dh) \end{aligned}$$

si $u \wedge v < \delta$, en donde el núcleo K' está dado por

$$K'f(u, v, w) = p_{u,v} \mathbb{E}(f(Qu, v, Qw)) + p_{v,u} \mathbb{E}(f(u, Qv, Qw)) + q_{u,v} \mathbb{E}(f(Qu, Qv, Qw)).$$

En otro caso tomamos $((X^x, X^y), Z)$ como un coupling independiente. Con esto, si $z \leq x \wedge y < \delta$ entonces $Z_t^z \leq X_t^x \wedge X_t^y$ en el intervalo $[0, \bar{\sigma}]$. Luego, si

$$\bar{\sigma}' = \inf\{t \geq 0 : Z_t^z \geq \delta\},$$

tendremos que $\mathbb{P}_{x,y,z}(\bar{\sigma} \leq \bar{\sigma}') = 1$ siempre que $z \leq x \wedge y$. Con esto, y usando el Lema 2.3

$$\begin{aligned} \mathbb{E}_{x,y}(e^{\gamma\bar{\sigma}}) &= \mathbb{E}_{x,y,x \wedge y}(e^{\gamma\bar{\sigma}}) \\ &\leq \mathbb{E}_{x,y,x \wedge y}(e^{\gamma\bar{\sigma}'}) \\ &= \mathbb{E}_{x \wedge y}(e^{\gamma\bar{\sigma}'}) \\ &\leq G_{\delta}^M(\gamma). \end{aligned}$$

□

Observación El tiempo de parada definido por

$$\tilde{\sigma} = \inf\{t \geq 0 : X_t \wedge Y_t \geq \hat{\varepsilon}\}$$

satisface $\mathbb{P}_{x,y}(\sigma \leq \tilde{\sigma}) = 1$. Repitiendo el argumento anterior con $\tilde{\sigma}$ en vez de $\bar{\sigma}$, entonces

$$\mathbb{E}_{x,y}(e^{\gamma\sigma}) \leq G_{\hat{\varepsilon}}^M(\gamma).$$

2.2. Minoración de algunas probabilidades

Para la demostración del resultado principal de este trabajo se necesitan varios Lemas.

Lema 2.6 *Sea (X^x, X^y) un coupling con generador dado por (2.6). Entonces*

$$\theta \doteq \inf_{(x,y) \in B_{\delta,\varepsilon}} \mathbb{P}_{x,y}(\bar{\sigma} \leq \sigma) \geq e^{-2\Lambda(\hat{\varepsilon},0)} > 0, \quad (2.13)$$

en donde $B_{\delta,\varepsilon} = [0, \varepsilon]^2 \setminus [\delta, \varepsilon]^2$.

Observación El conjunto $B_{\delta,\varepsilon}$ es simétrico, en el sentido que $(x, y) \in B_{\delta,\varepsilon}$ si y solamente si $(y, x) \in B_{\delta,\varepsilon}$. Además, de su definición es fácil ver que si (por ejemplo) $x > y$ entonces $(x, y) \in B_{\delta,\varepsilon}$ si y solamente si $x \leq \varepsilon$ e $y < \delta$.

DEMOSTRACIÓN. Sea $(x, y) \in B_{\delta,\varepsilon}$ y sean $(T_n^{x,y})_{n \geq 1}$ los tiempos de salto del par (X^x, X^y) . Por la observación anterior, podemos suponer que $x > y$. Dado que $\hat{\varepsilon} > \varepsilon + \delta$, en el evento $\{T_1^{x,y} > \delta - y\}$ se tiene que

$$X_{\delta-y}^x = x + \delta - y \leq \delta + \varepsilon < \hat{\varepsilon}, \quad X_{\delta-y}^y = y + \delta - y = \delta$$

por lo que se tiene la inclusión

$$\{T_1^{x,y} > \delta - y\} \subseteq \{\bar{\sigma} \leq \sigma\}.$$

Luego,

$$\begin{aligned} \mathbb{P}_{x,y}(\bar{\sigma} \leq \sigma) &\geq \mathbb{P}_{x,y}(T_1^{x,y} > \delta - y) \\ &= e^{-\int_0^{\delta-y} \beta(x+s, y+s) ds} \end{aligned}$$

De la desigualdad (2.7) se deduce que

$$\int_0^{\delta-y} \beta(x+s, y+s) ds \leq \Lambda(\delta-y, x) + \Lambda(\delta-y, y).$$

Aplicando la desigualdad (1.23) a ambos términos del miembro derecho se tiene que

$$\int_0^{\delta-y} \beta(x+s, y+s) ds \leq \Lambda(\delta+x-y, 0) + \Lambda(\delta, 0),$$

y dada la monotonía de Λ y el hecho de que $|x-y| \leq \varepsilon$, el primer término del lado derecho es a lo más $\Lambda(\delta+\varepsilon, 0)$. Por último, usando nuevamente la monotonía de Λ y la definición de $\hat{\varepsilon}$ se ve que

$$\mathbb{P}_{x,y}(\bar{\sigma} \leq \sigma) \geq e^{-2\Lambda(\hat{\varepsilon},0)}.$$

De esto es posible concluir, pues (x, y) es un elemento arbitrario de $B_{\delta,\varepsilon}$. \square

Lema 2.7 Sea $b \in [0, 1)$ tal que $H[0, b] > 0$, $x, y \geq \delta$ y $t_0 > \frac{b}{1-b}|x - y|$. Existe un coupling $(\tilde{X}^x, \tilde{X}^y)$ entre X^x y X^y , tal que la probabilidad $\mathbb{P}(\tilde{X}_s^x = \tilde{X}_s^y, \forall s \geq t_0)$ está minorada por

$$r_{t_0}(x, y) = \frac{1}{1 - \mu} \int_0^b \left(\int_{\frac{h}{1-h}|x-y|}^{t_0} e^{-\Lambda(s+x,0)} ds \right) e^{-\Lambda(h(x+t_0),0)} H(dh) \quad (2.14)$$

Más aún, si $A_{\delta, \hat{\varepsilon}} = \{(x, y) : \delta < x, y \leq \hat{\varepsilon}\} = [\delta, \hat{\varepsilon}]^2$ y $t_0 > \frac{b}{1-b}\hat{\varepsilon}$, entonces

$$\inf_{(x,y) \in A_{\delta, \hat{\varepsilon}}} r_{t_0}(x, y) \geq \frac{H[0, b]e^{-\Lambda(b(\hat{\varepsilon}+t_0),0)}}{1 - \mu} \int_{\frac{b}{1-b}\hat{\varepsilon}}^{t_0} e^{-\Lambda(s+\hat{\varepsilon},0)} ds > 0 \quad (2.15)$$

DEMOSTRACIÓN. Esta demostración está inspirada en el Lema 4.1 de [4]. Sin pérdida de generalidad, supongamos que $x > y$. Construimos el coupling $(\tilde{X}^x, \tilde{X}^y)$ a través de sus tiempos de salto (\tilde{T}_n^x) y (\tilde{T}_n^y) . Además, usando la misma v.a. Q para el primer salto de ambas componentes, condicionalmente a $Q = h$, el par $(\tilde{T}_1^x, \tilde{T}_1^y)$ se elige como el coupling maximal entre las v.a.

$$T_1^x \quad y \quad T_1^y + \frac{h}{1-h}(x - y), \quad (2.16)$$

donde T_1^w tiene densidad $f_w(r) = \lambda(w + r)e^{-\Lambda(r,w)}$, y escogemos

$$\tilde{X}_{\tilde{T}_1^x}^x = h(x + \tilde{T}_1^x), \quad \tilde{Y}_{\tilde{T}_1^y}^y = h(y + \tilde{T}_1^y).$$

El resto de los tiempos de salto $(\tilde{T}_n^x, \tilde{T}_n^y)_{n \geq 2}$ los escogemos independientes. Notemos que en el evento $C_h \doteq \left\{ \tilde{T}_2^y - \tilde{T}_1^y \geq \frac{h}{1-h}(x - y) \right\} \cap \left\{ \tilde{T}_1^x = \tilde{T}_1^y + \frac{h}{1-h}(x - y) \right\} \cap \{ \tilde{T}_1^y \leq t_0 \} \cap \{ Q = h \}$ se tiene

$$\begin{aligned} \tilde{X}_{\tilde{T}_1^x}^x &= h(x + \tilde{T}_1^x) \\ &= h \left(x + \tilde{T}_1^y + \frac{h}{1-h}(x - y) \right) \\ &= h \left(y + \tilde{T}_1^y + \frac{1}{1-h}(x - y) \right) \\ &= h(y + \tilde{T}_1^y) + \frac{h}{1-h}(x - y) \\ &= \tilde{X}_{\tilde{T}_1^y}^y + \tilde{T}_1^x - \tilde{T}_1^y \\ &= \tilde{X}_{\tilde{T}_1^x}^y \end{aligned}$$

y esto ocurre en el intervalo $[0, t_0]$. De esto obtenemos que

$$\begin{aligned} \mathbb{P}(\tilde{X}_s^x = \tilde{X}_s^y, \forall s \geq t_0) &\geq \int_0^1 \mathbb{P}(C_h) H(dh) \\ &\geq \int_0^a \mathbb{P}(C_h) H(dh) \end{aligned} \quad (2.17)$$

La probabilidad en (2.17) es igual a

$$\begin{aligned} &\mathbb{E} \left[\mathbf{1}_{\{\tilde{T}_1^x = \tilde{T}_1^y + \frac{h}{1-h}(x-y), \tilde{T}_1^y \leq t_0\}} \mathbb{E} \left(\mathbf{1}_{\{\tilde{T}_2^y - \tilde{T}_1^y \geq \frac{h}{1-h}(x-y)\}} \middle| (\tilde{T}_1^x, \tilde{T}_1^y) \right) \right] \\ &= \mathbb{E} \left[\mathbf{1}_{\{\tilde{T}_1^x = \tilde{T}_1^y + \frac{h}{1-h}(x-y), \tilde{T}_1^y \leq t_0\}} e^{-\Lambda\left(\frac{h}{1-h}(x-y), h(y + \tilde{T}_1^y)\right)} \right] \end{aligned}$$

2.3. RESULTADO PRINCIPAL

Notando que $\tilde{Y}_{\tilde{T}_1^y} \leq h(y + t_0) \leq h(x + t_0)$ en $\{\tilde{T}_1^y \leq t_0\}$, y usando la cota (1.23) se tiene que

$$\sup_{0 \leq w \leq h(y+t)} \Lambda \left(\frac{h}{1-h}(x-y), w \right) \leq \Lambda(h(x+t_0), 0)$$

de donde

$$\mathbb{P}(\tilde{X}_s^x = \tilde{X}_s^y, \forall s \geq t_0) \geq \int_0^b \mathbb{P} \left(\tilde{T}_1^x = \tilde{T}_1^y + \frac{h}{1-h}(x-y), \tilde{T}_1^y \leq t_0 \right) e^{-\Lambda(h(x+t_0), 0)} H(ds) \quad (2.18)$$

Ahora, si $0 < h < b$ tenemos que la igualdad entre las variables en (2.16) se alcanza antes del tiempo t_0 con probabilidad (maxima) igual a

$$I(t_0, x, y) = \int_{\frac{h}{1-h}(x-y)}^{t_0} f_x(s) \wedge f_y \left(s - \frac{h}{1-h}(x-y) \right) ds,$$

de acuerdo a lo discutido en la Subseccion 1.3.1.

Dado que $x > y > \delta$ tenemos que $f_w(r) \geq \eta e^{-\Lambda(r, w)}$ para todo $r > 0$ y $w = x, y$. De esta forma, $f_x(s) \geq \eta e^{-\Lambda(s+x, 0)}$. Ademas, utilizando la desigualdad (1.23) y la monotona de Λ se obtiene que

$$\Lambda \left(s - \frac{h}{1-h}(x-y), y \right) \leq \Lambda \left(s + x - \frac{x-y}{1-h}, 0 \right) \leq \Lambda(s+x, 0)$$

de forma que

$$\begin{aligned} f_y \left(s - \frac{h}{1-h}(x-y) \right) &\geq \lambda \left(y + s - \frac{h}{1-h}(x-y) \right) e^{-\Lambda(s+x, 0)} \\ &\geq \eta e^{-\Lambda(s+x, 0)}, \end{aligned}$$

por lo que Por ultimo, notando que $\eta > \frac{1}{1-\mu}$ obtenemos que

$$I(t_0, x, y) \geq \frac{1}{1-\mu} \int_{\frac{h}{1-h}(x-y)}^{t_0} e^{-\Lambda(s+x, 0)} ds \quad (2.19)$$

La desigualdad (2.15) se obtiene reemplazando (2.19) en (2.18), y la estimacion (2.15) es directa. \square

2.3. Resultado Principal

Como puede apreciarse en el Lema 2.7, es interesante estudiar los tiempos de entrada del par (X^x, X^y) al conjunto $A_{\delta, \varepsilon}$, ya que una vez que $(X_t^x, X_t^y) \in A_{\delta, \varepsilon}$ se puede controlar la probabilidad de que en algun tiempo $t_0 > 0$ se produzca el acoplamiento en el intervalo $[t, t + t_0]$, a saber, esta probabilidad esta minorada por el nfimo en (2.15). Lamentablemente, el tiempo t en el argumento anterior no sera fijo, sino aleatorio. Aun ası, una aplicacion de la propiedad de Markov fuerte permite hacer valido el razonamiento incluso en ese caso.

Específicamente, sea

$$T(s) = \inf\{t \geq 0 : (X_{t+s}^x, X_{t+s}^y) \in A_{\delta, \hat{\varepsilon}}\} \quad (2.20)$$

el tiempo de entrada al conjunto $A_{\delta, \hat{\varepsilon}}$ medido desde s y $T = T(0)$. A priori, T no es fácil de controlar, pero puede ser descompuesto en términos de los tiempos τ , σ y $\bar{\sigma}$, definidos en (2.8)–(2.10), de la siguiente forma.

Sea $R_0 \equiv 0$. Para $k \in \mathbb{N}_0$ se definen, inductivamente, los tiempos

$$\tau_{k+1} = \tau(R_k) \quad (2.21)$$

$$\sigma_{k+1} = \sigma(R_k + \tau_{k+1}) \quad (2.22)$$

$$\bar{\sigma}_{k+1} = \bar{\sigma}(R_k + \tau_{k+1}) \quad (2.23)$$

$$W_{k+1} = \tau_{k+1} + \sigma_{k+1} \wedge \bar{\sigma}_{k+1} \quad (2.24)$$

$$R_{k+1} = R_k + W_{k+1}. \quad (2.25)$$

La idea detrás de estas definiciones es descomponer los intentos de entrada a $A_{\delta, \hat{\varepsilon}}$ en dos fases. En primer lugar, se espera un tiempo τ , en el cual se sabe que $(X_\tau^x, X_\tau^y) \in [0, \varepsilon]^2$. A continuación, se deja correr el proceso hasta el tiempo $\sigma \wedge \bar{\sigma}$ y diremos que se ha realizado un *intento de entrada a $A_{\delta, \hat{\varepsilon}}$* . Si resulta que $\bar{\sigma} \leq \sigma$ entonces $(X_{\tau+\sigma \wedge \bar{\sigma}}^x, X_{\tau+\sigma \wedge \bar{\sigma}}^y) \in A_{\delta, \hat{\varepsilon}}$, se obtiene que $T = W_1$ y se termina. En caso contrario, se ha fallado un intento y se vuelve a reiniciar el procedimiento desde la posición final. Resulta entonces que W_k es la duración del k -ésimo intento de entrada, y R_k es el tiempo transcurrido desde 0 hasta el final de dicho intento.

Con esto, es fácil ver que si

$$N = \inf\{k \in \mathbb{N} : \bar{\sigma}_k \leq \sigma_k\}$$

entonces se tiene que

$$T = \sum_{j=1}^N W_j = R_N.$$

Proposición 2.8 Sean $x_0 > \hat{\varepsilon}$, $p > 1$, θ definido en la ecuación (2.13), $p^{-1} + q^{-1} = 1$ y

$$\psi_p^{x_0}(\gamma) \doteq \frac{1}{p} \ln(G_{\hat{\varepsilon}}^M(p\gamma)) + \ln(1 + \gamma x_0)$$

con $G_{\hat{\varepsilon}}^M(\gamma)$ definido en (2.5). Sean además

$$\tilde{\theta} \doteq 1 - (1 - \theta)^{1/q} \quad (2.26)$$

y

$$\tilde{\gamma} \doteq \sup\{\gamma > 0 : \psi_p^{x_0}(\gamma) < -\ln(1 - \tilde{\theta})\}, \quad (2.27)$$

Si se define

$$\hat{\gamma} \doteq \min\left\{\frac{1}{\varepsilon}(\eta(1 - \mu)\varepsilon - 2), -\frac{1}{p\delta} \ln(1 - e^{-M\delta}), \tilde{\gamma}\right\}. \quad (2.28)$$

se tiene la estimación

$$\sup_{(x,y) \in [0, x_0]^2} \mathbb{E}_{x,y}(e^{\gamma T}) \leq \frac{\tilde{\theta} e^{\psi_p^{x_0}(\gamma)}}{1 - e^{\psi_p^{x_0}(\gamma)}(1 - \tilde{\theta})}, \quad (2.29)$$

2.3. RESULTADO PRINCIPAL

para todo $\gamma \in (0, \hat{\gamma})$. Más aún,

$$\tilde{\gamma} \geq \frac{1}{2px_0} \ln \left(\frac{(1-\theta)^{1-p}}{e^{-M\hat{\varepsilon}} + (1-\theta)^{1-p}(1-e^{-M\hat{\varepsilon}})} \right). \quad (2.30)$$

DEMOSTRACIÓN. Dado que (R_k) es una sucesión creciente, la familia $\mathcal{G}_k = \mathcal{F}_{R_k}$ es una filtración. La familia de eventos $A_k = \{\bar{\sigma}_k \leq \sigma_k\}$ es adaptada a la filtración \mathcal{G}_k . En efecto,

$$A_k = \{(X_{R_k}^x, X_{R_k}^y) \in A_{\delta, \hat{\varepsilon}}\} \in \mathcal{G}_k.$$

Además, la sucesión (W_k) es adaptada a \mathcal{G}_k pues $W_k = R_k - R_{k-1}$. Por otra parte, si para $k \in \mathbb{N}_0$ se define $\mathcal{G}_{k+\frac{1}{2}} = \mathcal{F}_{R_k+\tau_{k+1}}$ entonces se tienen las inclusiones

$$\mathcal{G}_k \subset \mathcal{G}_{k+\frac{1}{2}} \subset \mathcal{G}_{k+1},$$

y τ_{k+1} es $\mathcal{G}_{k+\frac{1}{2}}$ -medible, para todo $k \in \mathbb{N}_0$.

Dados $(x, y) \in [0, x_0]^2$, $0 < \gamma < \hat{\gamma}$ y $k \in \mathbb{N}$ se tiene que

$$\begin{aligned} \mathbb{E}_{x,y}(e^{\gamma W_k} \mid \mathcal{G}_{k-1}) &= \mathbb{E}_{x,y}(e^{\gamma(\tau_k + \bar{\sigma}_k \wedge \sigma_k)} \mid \mathcal{G}_{k-1}) \\ &= \mathbb{E}_{x,y} \left(e^{\gamma \tau_k} \mathbb{E}_{x,y} \left(e^{\gamma(\bar{\sigma}_k \wedge \sigma_k)} \mid \mathcal{G}_{k-\frac{1}{2}} \right) \mid \mathcal{G}_k \right). \end{aligned}$$

Ahora, usando la propiedad de Markov fuerte sobre la esperanza condicional interior, se tiene que

$$\begin{aligned} \mathbb{E}_{x,y} \left(e^{\gamma(\bar{\sigma}_k \wedge \sigma_k)} \mid \mathcal{G}_{k-\frac{1}{2}} \right) &= \mathbb{E}_{X_{R_{k-1}+\tau_k}^x, X_{R_{k-1}+\tau_k}^y} (e^{\gamma(\bar{\sigma} \wedge \sigma)}) \\ &\leq \mathbb{E}_{X_{R_{k-1}+\tau_k}^x, X_{R_{k-1}+\tau_k}^y} (e^{\gamma \sigma}) \end{aligned}$$

Por definición, $(X_{R_{k-1}+\tau_k}^x, X_{R_{k-1}+\tau_k}^y) \in [0, \varepsilon]^2$ y así el último término es a lo más

$$\sup_{(x,y) \in [0, \varepsilon]^2} \mathbb{E}_{x,y}(e^{\gamma \sigma}) \leq G_{\hat{\varepsilon}}^M(\gamma),$$

en donde se ha usado el Lema 2.5 y la observación que lo sucede. Esta aplicación es posible pues la elección de $\hat{\gamma}$ asegura que se satisfacen las hipótesis.

Reemplazando esto en (2.31) se obtiene la estimación

$$\mathbb{E}_{x,y}(e^{\gamma W_k} \mid \mathcal{G}_{k-1}) \leq G_{\hat{\varepsilon}}^M(\gamma) \mathbb{E}_{x,y}(e^{\gamma \tau_k} \mid \mathcal{G}_{k-1}). \quad (2.31)$$

Nuevamente gracias a la propiedad de Markov fuerte la esperanza condicional del lado derecho es igual a

$$\mathbb{E}_{X_{R_{k-1}}^x, X_{R_{k-1}}^y} (e^{\gamma \tau}) \leq (1 + \gamma(X_{R_{k-1}}^x \vee X_{R_{k-1}}^y)),$$

en donde la desigualdad proviene del Lema 2.4. Pero si $k > 1$ entonces

$$(X_{R_{k-1}}^x, X_{R_{k-1}}^y) \in [0, \hat{\varepsilon}]^2 \subset [0, x_0]^2,$$

y por hipótesis $(X_{R_0}^x, X_{R_0}^y) = (x, y) \in [0, x_0]^2$. En cualquier caso,

$$X_{R_{k-1}}^x \vee X_{R_{k-1}}^y \leq x_0$$

de modo que reemplazando esto en (2.31) se obtiene que

$$\mathbb{E}_{x,y}(e^{\gamma W_k} \mid \mathcal{G}_{k-1}) \leq G_{\hat{\varepsilon}}^M(\gamma)(1 + \gamma x_0) = e^{\psi_1^{x_0}(\gamma)}. \quad (2.32)$$

Por otro lado, dada la definición del evento A_k se tiene que

$$\begin{aligned} \mathbb{E}_{x,y}(e^{\gamma W_k} \mathbf{1}_{A_k^c} \mid \mathcal{G}_{k-1}) &= \mathbb{E}_{x,y} \left(e^{\gamma \tau_k} \mathbb{E}_{x,y} \left(e^{\gamma(\bar{\sigma}_k \wedge \sigma_k)} \mathbf{1}_{A_k^c} \mid \mathcal{G}_{k-\frac{1}{2}} \right) \mid \mathcal{G}_{k-1} \right) \\ &= \mathbb{E}_{x,y} \left(e^{\gamma \tau_k} \mathbb{E}_{x,y} \left(e^{\gamma \sigma_k} \mathbf{1}_{A_k^c} \mid \mathcal{G}_{k-\frac{1}{2}} \right) \mid \mathcal{G}_{k-1} \right) \end{aligned} \quad (2.33)$$

Aplicando la desigualdad de Hölder a la esperanza condicional interior, y repitiendo un argumento anterior se tiene que

$$\begin{aligned} \mathbb{E}_{x,y} \left(e^{\gamma \sigma_k} \mathbf{1}_{A_k^c} \mid \mathcal{G}_{k-\frac{1}{2}} \right) &= \mathbb{E}_{X_{R_{k-1}+\tau_k}^x, X_{R_{k-1}+\tau_k}^y} (e^{\gamma \sigma} \mathbf{1}_{A_1^c}) \\ &\leq \left[\mathbb{E}_{X_{R_{k-1}+\tau_k}^x, X_{R_{k-1}+\tau_k}^y} (e^{p\gamma \sigma}) \right]^{1/p} \mathbb{P}_{X_{R_{k-1}+\tau_k}^x, X_{R_{k-1}+\tau_k}^y} (\sigma \leq \bar{\sigma})^{1/q} \\ &\leq [G_{\hat{\varepsilon}}^M(p\gamma)]^{1/p} \left[\sup_{(x,y) \in [0,\varepsilon]^2} \mathbb{P}_{x,y}(\sigma \leq \bar{\sigma}) \right]^{1/q}. \end{aligned} \quad (2.34)$$

Pero si $\theta > 0$ es la probabilidad definida en (2.13), entonces

$$\begin{aligned} \left[\sup_{(x,y) \in [0,\varepsilon]^2} \mathbb{P}_{x,y}(\sigma \leq \bar{\sigma}) \right]^{1/q} &= \left[1 - \inf_{(x,y) \in [0,\varepsilon]^2} \mathbb{P}_{x,y}(\bar{\sigma} \leq \sigma) \right]^{1/q} \\ &= [1 - \theta]^{1/q} \\ &= 1 - \tilde{\theta}, \end{aligned}$$

donde $\tilde{\theta}$ está definido en (2.26). Reemplazando primero en (2.34) y luego en (2.33) se concluye que

$$\mathbb{E}_{x,y}(e^{\gamma W_k} \mathbf{1}_{A_k^c} \mid \mathcal{G}_{k-1}) \leq [G_{\hat{\varepsilon}}^M(p\gamma)]^{1/p} (1 - \tilde{\theta}) \mathbb{E}_{x,y}(e^{\gamma \tau_k} \mid \mathcal{G}_{k-1}). \quad (2.35)$$

Repitiendo un argumento anterior se deduce que

$$\mathbb{E}_{x,y}(e^{\gamma W_k} \mathbf{1}_{A_k^c} \mid \mathcal{G}_{k-1}) \leq [G_{\hat{\varepsilon}}^M(p\gamma)]^{1/p} (1 - \tilde{\theta}) (1 + \gamma x_0) = (1 - \tilde{\theta}) e^{\psi_p^{x_0}(\gamma)}.$$

Para finalizar, veamos que la función $p \mapsto \psi_p^{x_0}(\gamma)$ es creciente. Para esto, basta ver que la función $h(p) = [G_{\hat{\varepsilon}}^M(p\gamma)]^{1/p}$ lo es. Como $G_{\hat{\varepsilon}}^M(\cdot)$ es el momento exponencial de una v.a. entonces es creciente. Por regla de la cadena, se tiene que

$$h'(p) = \frac{\gamma}{p} [G_{\hat{\varepsilon}}^M(p\gamma)]^{1/p-1} (G_{\hat{\varepsilon}}^M)'(p\gamma) > 0.$$

Así, es posible reemplazar $\psi_1^{x_0}$ por $\psi_p^{x_0}$ en (2.32), y la elección de $\hat{\gamma}$ asegura que puede aplicarse el Lema 2.2. Por otra parte,

$$\begin{aligned} e^{\psi_p^{x_0}(\gamma)} &= [G_{\hat{\varepsilon}}^M(p\gamma)]^{1/p} (1 + \gamma x_0) \\ &\leq [G_{\hat{\varepsilon}}^M(p\gamma)]^{1/p} e^{\gamma x_0} \\ &= \left[\frac{e^{(p\gamma - M)\hat{\varepsilon}} e^{p\gamma x_0}}{1 - e^{p\gamma \hat{\varepsilon}} (1 - e^{-M\hat{\varepsilon}})} \right]^{1/p} \end{aligned}$$

2.3. RESULTADO PRINCIPAL

según (2.5). Luego, $\psi_p^{x_0}(\gamma) < -\ln(1 - \tilde{\theta})$ si y solamente si $e^{\psi_p^{x_0}(\gamma)} < (1 - \tilde{\theta})^{-1}$, y esto último ocurre si

$$\frac{e^{p\gamma(\hat{\varepsilon}+x_0)}e^{-M\hat{\varepsilon}}}{1 - e^{p\gamma\hat{\varepsilon}}(1 - e^{-M\hat{\varepsilon}})} < (1 - \tilde{\theta})^{-p} = (1 - \theta)^{1-p},$$

o de forma equivalente, si

$$e^{p\gamma(\hat{\varepsilon}+x_0)}e^{-M\hat{\varepsilon}} + (1 - \theta)^{1-p}e^{p\gamma\hat{\varepsilon}}(1 - e^{-M\hat{\varepsilon}}) < (1 - \theta)^{1-p}.$$

Pero como $e^{p\gamma\hat{\varepsilon}} \leq e^{p\gamma(\hat{\varepsilon}+x_0)} \leq e^{2p\gamma x_0}$, para que se satisfaga esta desigualdad basta que

$$2p\gamma x_0 < \ln \left(\frac{(1 - \theta)^{1-p}}{e^{-M\hat{\varepsilon}} + (1 - \theta)^{1-p}(1 - e^{-M\hat{\varepsilon}})} \right).$$

De esta forma se ha probado la desigualdad (2.30). \square

A continuación, se construirá un coupling (\hat{X}^x, \hat{X}^y) tal que el tiempo

$$T^* = \inf\{t \geq 0 : \hat{X}_t^x = \hat{X}_t^y\} \quad (2.36)$$

tenga momento exponencial finito.

Teorema 2.9 *Dados $p > 1$, $b \in [0, 1)$ tal que $H[0, b] > 0$ y $t_0 > \frac{b}{1-b}\hat{\varepsilon}$, existe un coupling (\hat{X}^x, \hat{X}^y) tal que el tiempo T^* definido en (2.36) satisface*

$$\sup_{(x,y) \in [0, \varepsilon]^2} \mathbb{E}_{x,y}(e^{\gamma T^*}) \leq \frac{\hat{\theta} e^{\Psi_p^{t_0}(\gamma)}}{1 - e^{\Psi_p^{t_0}(\gamma)}(1 - \hat{\theta})}, \quad (2.37)$$

para $0 < \gamma < \hat{\gamma}$, en donde $\hat{\gamma}$ está definido en (2.28). Aquí, $\hat{\theta}$ es el ínfimo en (2.15), esto es,

$$\hat{\theta} \doteq \inf_{(x,y) \in A_{\delta, \varepsilon}} r_{t_0}(x, y) > 0$$

y

$$\Psi_p^{t_0}(\gamma) \doteq \gamma t_0 + \psi_p^{\hat{\varepsilon}+t_0}(\gamma).$$

DEMOSTRACIÓN. La idea de la demostración es iterar el procedimiento que precede a la Proposición 2.8 junto con el coupling construido en el Lema 2.7. Fijemos $t_0 > \frac{b}{1-b}\hat{\varepsilon}$.

El proceso (\hat{X}^x, \hat{X}^y) será definido inductivamente como sigue. Sea $R_0 \equiv 0$. Partiendo de (x, y) tomamos (\hat{X}^x, \hat{X}^y) con generador como en la ecuación (2.6) hasta el tiempo $T_1 = T(R_0) = T$ definido en (2.20), esto es, $(\hat{X}_t^x, \hat{X}_t^y)$ sigue la dinámica dada por (2.6) para $t \in [0, T_1]$. Luego, se toma (\hat{X}^x, \hat{X}^y) como el coupling $(\tilde{X}^x, \tilde{X}^y)$ del Lema 2.7 en el intervalo $[T_1, T_1+t_0]$, partiendo de $(\hat{X}_{T_1}^x, \hat{X}_{T_1}^y)$, es decir, $(\hat{X}_t^x, \hat{X}_t^y) = \left(\tilde{X}_t^{\hat{X}_{T_1}^x}, \tilde{X}_t^{\hat{X}_{T_1}^y} \right)$ para $t \in [T_1, T_1+t_0]$.

Notemos que por definición $(\hat{X}_{T_1}^x, \hat{X}_{T_1}^y) \in [\delta, \hat{\varepsilon}]^2$ por lo que todo está bien definido. Diremos que en el intervalo $[0, T_1 + t_0]$ se ha completado un ciclo, y se define $R_1 = W_1 = T_1 + t_0$, la duración del primer ciclo.

Ahora, se rellena el proceso de forma inductiva para $t > R_1$. Específicamente, dado $n \in \mathbb{N}$, si (\hat{X}^x, \hat{X}^y) ya está definido en $[0, R_n]$, escogemos $(\hat{X}_t^x, \hat{X}_t^y)$ como cualquier coupling Markoviano con generador (2.6) hasta el tiempo $T_{n+1} = T(R_n)$ y como $\left(\tilde{X}_t^{\hat{X}_{R_n}^x}, \tilde{X}_t^{\hat{X}_{R_n}^y} \right)$ para

$t \in [R_n + T_{n+1}, R_n + T_{n+1} + t_0]$. Finalmente, $W_{n+1} = T_{n+1} + t_0$ es la duración del ciclo $(n+1)$ -ésimo y $R_{n+1} = R_n + W_{n+1}$ es el tiempo real transcurrido hasta el final de dicho ciclo.

Al igual que antes, si

$$M = \inf\{n \in \mathbb{N} : \hat{X}_{R_n}^x = \hat{X}_{R_n}^y\},$$

se tiene la estimación

$$T^* \leq \sum_{j=1}^M W_j = R_M.$$

Consideremos la filtración $\mathcal{G}_n = \mathcal{F}_{R_n}$ y sea $\mathcal{G}_{n+\frac{1}{2}} = \mathcal{F}_{R_n+T_{n+1}}$. Dados $\gamma < \hat{\gamma}$, $n \in \mathbb{N}$, $p > 1$ y $(x, y) \in [0, \varepsilon]^2$ se tiene, gracias a la propiedad de Markov fuerte, que

$$\begin{aligned} \mathbb{E}_{x,y}(e^{\gamma W_n} \mid \mathcal{G}_{n-1}) &= e^{\gamma t_0} \mathbb{E}_{x,y}(e^{\gamma T_n} \mid \mathcal{G}_{n-1}) \\ &= e^{\gamma t_0} \mathbb{E}_{\hat{X}_{R_{n-1}}^x, \hat{X}_{R_{n-1}}^y}(e^{\gamma T}) \end{aligned}$$

De la Proposición 2.8 se sigue que

$$\mathbb{E}_{x,y}(e^{\gamma W_n} \mid \mathcal{G}_{n-1}) \leq e^{\gamma t_0 + \psi_p^{\varepsilon+t_0}(\gamma)}, \quad (2.38)$$

pues $(\hat{X}_{R_n}^x, \hat{X}_{R_n}^y) \in [0, \hat{\varepsilon} + t_0]^2$ para todo $n \in \mathbb{N}_0$.

Por otra parte, si se define $A_n \doteq \{\hat{X}_{R_n}^x = \hat{X}_{R_n}^y\} \in \mathcal{G}_n$ entonces

$$\begin{aligned} \mathbb{E}_{x,y}(e^{\gamma W_n} \mathbf{1}_{A_n^c} \mid \mathcal{G}_{n-1}) &= e^{\gamma t_0} \mathbb{E}_{x,y}(e^{\gamma T_n} \mathbf{1}_{A_n^c} \mid \mathcal{G}_{n-1}) \\ &= e^{\gamma t_0} \mathbb{E}_{x,y}(e^{\gamma T_n} \mathbb{P}_{x,y}(A_n^c \mid \mathcal{G}_{n-\frac{1}{2}}) \mid \mathcal{G}_{n-1}) \end{aligned} \quad (2.39)$$

Pero, en virtud de la Propiedad de Markov fuerte

$$\begin{aligned} \mathbb{P}_{x,y}(A_n^c \mid \mathcal{G}_{n-\frac{1}{2}}) &= \mathbb{P}_{\hat{X}_{R_{n-1}+T_n}^x, \hat{X}_{R_{n-1}+T_n}^y}(\tilde{X}_{t_0}^x \neq \tilde{X}_{t_0}^y) \\ &\leq \sup_{(x,y) \in A_{\delta,\varepsilon}} \mathbb{P}_{x,y}(\tilde{X}_{t_0}^x \neq \tilde{X}_{t_0}^y) \end{aligned} \quad (2.40)$$

Usando el Lema 2.7 se obtiene que

$$\sup_{(x,y) \in A_{\delta,\varepsilon}} \mathbb{P}_{x,y}(\tilde{X}_{t_0}^x \neq \tilde{X}_{t_0}^y) \leq 1 - \inf_{(x,y) \in A_{\delta,\varepsilon}} r_{t_0}(x, y) = 1 - \hat{\theta}.$$

Reemplazando esto en (2.40) y luego en (2.39) se obtiene que

$$\begin{aligned} \mathbb{E}_{x,y}(e^{\gamma W_n} \mathbf{1}_{A_n^c} \mid \mathcal{G}_{n-1}) &\leq (1 - \hat{\theta}) e^{\gamma t_0} \mathbb{E}_{x,y}(e^{\gamma T_n} \mid \mathcal{G}_{n-1}) \\ &\leq (1 - \hat{\theta}) e^{\gamma t_0 + \psi_p^{\varepsilon+t_0}(\gamma)} \end{aligned}$$

en donde se ha usado la estimación en (2.38). Con esto, se puede aplicar el Lema 2.2 con $x_0 = \hat{\varepsilon} + t_0$ para deducir (2.37). \square

En la Ilustración 2.1 se aprecia una trayectoria exitosa del coupling descrito hasta ahora.

Este Teorema da un control sobre el momento exponencial del tiempo de parada T^* en el caso en que la condición inicial (x, y) se encuentra en el compacto $[0, \varepsilon]^2$. Usando el tiempo τ definido en (2.8) es fácil extender el argumento precedente a condiciones iniciales arbitrarias.

2.3. RESULTADO PRINCIPAL

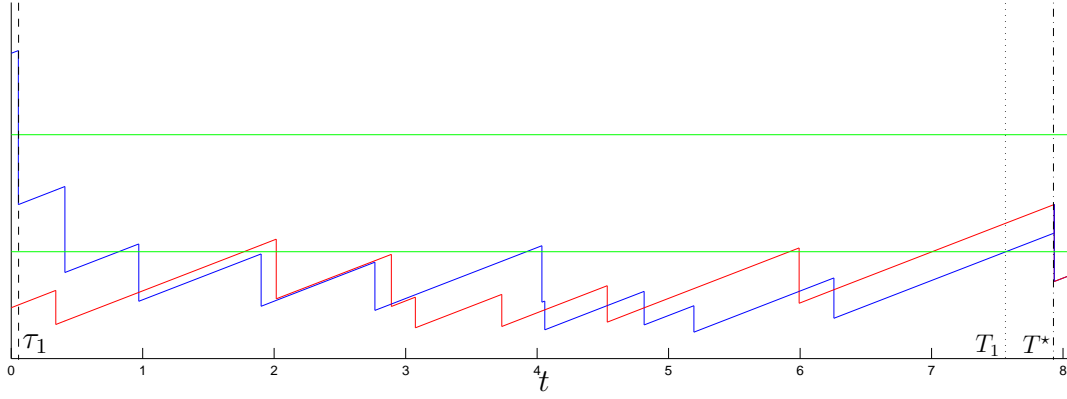


Ilustración 2.1: Trayectoria exitosa del coupling (\hat{X}^x, \hat{X}^y) con $x = 6, y = 2$.

Corolario 2.10 *Bajo las mismas hipótesis del Teorema 2.9, se tiene que*

$$\mathbb{E}_{x,y}(e^{\gamma T^*}) \leq (1 + \gamma(x \vee y)) \sup_{(u,v) \in [0,\varepsilon]^2} \mathbb{E}_{u,v}(e^{\gamma T^*}) \quad (2.41)$$

$$\leq \frac{\tilde{\theta} e^{\psi_p^{x_0}(\gamma)}}{1 - e^{\psi_p^{x_0}(\gamma)}(1 - \tilde{\theta})} \quad (2.42)$$

para todo $(x, y) \in \mathbb{R}_+^2$ y $\gamma < \hat{\gamma}$.

DEMOSTRACIÓN. Consideremos el coupling (\hat{X}^x, \hat{X}^y) y sea τ el tiempo de parada definido en (2.8). Sea además S^* el tiempo de parada definido por

$$S^* = \inf\{s \geq 0 : \hat{X}_{\tau+s}^x = \hat{X}_{\tau+s}^y\}.$$

Notemos que $T^* = \tau + S^*$. Usando la propiedad de Markov fuerte se obtiene que, para cualquier $(x, y) \in \mathbb{R}_+^2$

$$\begin{aligned} \mathbb{E}_{x,y}(e^{\gamma T^*}) &= \mathbb{E}_{x,y}(e^{\gamma\tau + \gamma S^*}) \\ &= \mathbb{E}_{x,y}(e^{\gamma\tau} \mathbb{E}_{x,y}(e^{\gamma S^*} \mid \mathcal{F}_\tau)). \end{aligned} \quad (2.43)$$

Gracias a que $(\hat{X}_\tau^x, \hat{X}_\tau^y) \in [0, \varepsilon]^2$ se tiene, usando el Teorema 2.9, que

$$\begin{aligned} \mathbb{E}_{x,y}(e^{\gamma S^*} \mid \mathcal{F}_\tau) &= \mathbb{E}_{\hat{X}_\tau^x, \hat{X}_\tau^y}(e^{\gamma T^*}) \\ &\leq \sup_{(u,v) \in [0,\varepsilon]^2} \mathbb{E}_{u,v}(e^{\gamma T^*}). \end{aligned}$$

Reemplazando esto en (2.43) y usando el Lema 2.4 se obtiene (2.41). \square

Observación En particular, se deduce que para todo $x, y \geq 0$, el conjunto $[\delta, \hat{\varepsilon} + t_0]^2$ es regenerativo para el par (\hat{X}^x, \hat{X}^y) , con largo de ciclo dado por T . Luego, el proceso original X tiene una medida invariante y como T tiene momentos exponenciales, dicha medida invariante tiene momentos finitos.

Usando el Lema 1.20, se deduce el siguiente

Teorema 2.11 *Dados $x, y \in \mathbb{R}_+$, $p > 1$, $b \in [0, 1)$ tal que $H[0, b] > 0$ y $t_0 > \frac{b}{1-b}\hat{\varepsilon}$ se tiene que*

$$\|P_t(x, \cdot) - P_t(y, \cdot)\|_{VT} \leq (1 + \gamma(x \vee y)) \frac{\hat{\theta} e^{\Psi_p^{t_0}(\gamma)}}{1 - e^{\Psi_p^{t_0}(\gamma)}(1 - \hat{\theta})} e^{-\gamma t} \quad (2.44)$$

para todo $0 < \gamma < \hat{\gamma}$. En particular, si μ y ν tienen primer momento finito, existe una constante $C = C(\mu, \nu, \gamma, t_0, p)$ tal que

$$\|\mu P_t - \nu P_t\|_{VT} \leq C e^{-\gamma t}, \quad (2.45)$$

en donde

$$C(\mu, \nu, \gamma, t_0, p) = (1 + \gamma(\mu_1 + \nu_1)) \frac{\hat{\theta} e^{\Psi_p^{t_0}(\gamma)}}{1 - e^{\Psi_p^{t_0}(\gamma)}(1 - \hat{\theta})}$$

y

$$\mu_1 = \int_0^\infty x \mu(dx), \quad \nu_1 = \int_0^\infty x \nu(dx).$$

DEMOSTRACIÓN. La estimación (2.44) proviene de aplicar el Corolario 2.10 y la desigualdad (1.12). La desigualdad (2.45) se concluye de (2.44) y el Lema 1.20. \square

Observación Es importante observar que las constantes que aparecen en (2.44) y (2.45) contienen el término $e^{\gamma t_0}$, por lo que los lados derechos de estas desigualdades son mayores a 1 si $t < t_0$. Por consiguiente, dichas estimaciones sólo aportan información para $t \geq t_0$.

2.4. Aplicación

En esta última sección se aplicarán los resultados obtenidos al proceso TCP. Recordemos que este proceso pertenece a la clase AIMD y tiene generador dado por

$$Af(x) = f'(x) + x \left(f\left(\frac{x}{2}\right) - f(x) \right).$$

De aquí se puede identificar inmediatamente los parámetros $\lambda(x) = x$ y $H(dh) = \delta_{1/2}(dh)$. Con esto, se tiene que $\mu = \frac{1}{2}$. Dada la forma de la función λ , para que se satisfaga la Hipótesis 1 basta con escoger $\delta = 2 + \alpha$ con $\alpha > 0$, y en ese caso

$$\eta = \delta = 2 + \alpha > 2 = \frac{1}{1 - \mu}.$$

Además, para esta elección de parámetros, tenemos que

$$2\bar{\eta} = 2 < \delta$$

por lo que basta escoger $\varepsilon > \delta$. En particular, se tomará $\varepsilon = 2(1 + \alpha)$. Por último, dado que $\hat{\varepsilon} > \varepsilon + \delta = 4 + 3\alpha$ se escogerá $\hat{\varepsilon} = 4(1 + \alpha)$. Ahora bien, por definición se tiene que

$$M' = \hat{\varepsilon} = 4(1 + \alpha)$$

2.4. APLICACIÓN

y luego $M = 2M' = 8(1 + \alpha)$ por lo que la Hipótesis 2 se satisface.

Por otra parte, dado que $H[0, 1/2] = 1$, basta escoger $b = 1/2$ con lo que $t_0 > \hat{\varepsilon} = 4(1 + \alpha)$. Para obtener una estimación de la tasa de convergencia $\hat{\gamma}$ expuesta en (2.28), primero se observa que la función

$$\begin{aligned} g(\alpha) &= \min \left\{ \frac{1}{\varepsilon}(\eta(1 - \mu)\varepsilon - 2), -\frac{1}{p\delta} \ln(1 - e^{-M\delta}) \right\} \\ &= \min \left\{ \frac{\alpha}{1 + \alpha}, -\frac{1}{p(2 + \alpha)} \ln(1 - e^{-8(1+\alpha)(2+\alpha)}) \right\} \end{aligned}$$

definida en \mathbb{R}_+ tiene un máximo en el valor $\tilde{\alpha} = \tilde{\alpha}(p)$, solución de la ecuación

$$\frac{\alpha}{1 + \alpha} = -\frac{1}{p(2 + \alpha)} \ln(1 - e^{-8(1+\alpha)(2+\alpha)}).$$

Es directo observar que $p \mapsto \tilde{\alpha}(p)$ es una función decreciente (por medio del teorema de la función inversa) y luego, escogemos $p = 1 + \zeta$ con $\zeta > 0$. Numéricamente se encuentra que $\tilde{\alpha}(1 + \zeta) \sim 5,62 \times 10^{-8}$ para ζ pequeño.

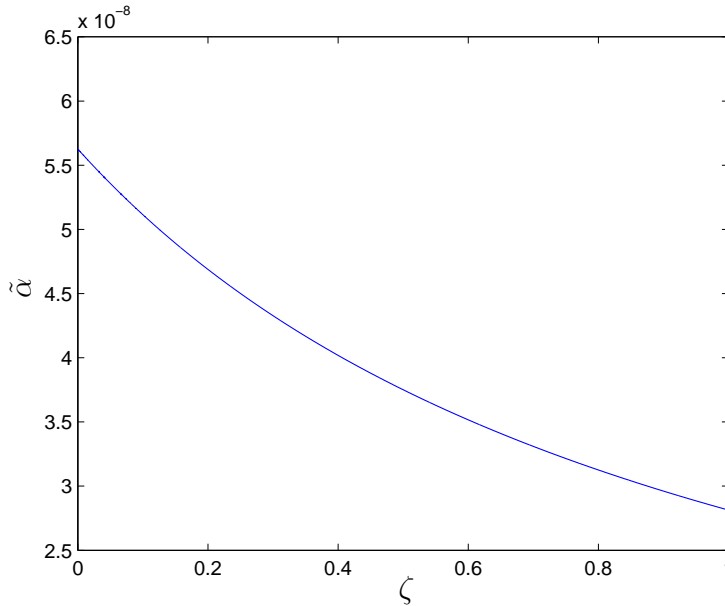


Ilustración 2.2: Comportamiento de la solución $\tilde{\alpha}$ respecto de $p = 1 + \zeta$.

En la Ilustración 2.2 se puede apreciar más precisamente el comportamiento de esta solución, para valores de ζ cercanos a 0. Esto ya provee una primera evidencia de cómo hay que escoger los parámetros en este caso: p debe ser cercano a 1, pero la distancia entre δ y ε (igual a $\tilde{\alpha}$) debe ser escogida cuidadosamente para maximizar la tasa de convergencia.

Obviamente, el valor obtenido aquí es estimativo y dar una cantidad más precisa requiere el cálculo explícito del resto de las constantes que aparecen en el Teorema 2.11. Sin embargo, los resultados numéricos ya dan cuenta de la gran diferencia (de 2 órdenes de magnitud aproximadamente) que se produce en las tasas encontradas en [4].

Conclusión

El principal resultado probado en este trabajo de memoria, el Teorema 2.11, establece condiciones suficientes para una convergencia exponencial al equilibrio de los procesos de la clase AIMD. Además de establecer dichas condiciones, se da una expresión explícita para la tasa a la cual ocurre esta convergencia. Como subproducto de este análisis se prueba, bajo las mismas condiciones, la existencia de una medida invariante la cual tiene momentos de todos los órdenes. Específicamente, en el caso de la vida actual asociada a un proceso de renovación el Teorema 2.11 es una extensión del Teorema 1.25, el que sólo prueba la existencia de la tasa de convergencia pero no provee una forma explícita.

Aún así, en el caso del proceso TCP, la tasa obtenida es peor. Esto se debe a que la técnica utilizada durante el desarrollo de este trabajo es de un carácter más general y por ende aplica a más procesos, los que pueden tener un comportamiento peor que el ya citado TCP. Sin embargo, es este mismo punto el que potencia el carácter del resultado principal, pues provee un marco unificador en la clase AIMD. Dentro de la bibliografía revisada no se encuentran resultados de este tipo, sino más bien aplicaciones a casos particulares. En este sentido, los resultados obtenidos en este trabajo son de una generalidad que resulta novedosa en este contexto.

Quedan entonces abierta la cuestión de extender aún más estos resultados. Precisamente, una de las primeras generalizaciones posibles se encuentra en la dinámica determinista. Esto implica generar nuevas ideas, ya fuera de la clase AIMD; esto es, queda puesta la pregunta de si estos resultados siguen siendo válidos en la clase PDMP o en otra subclase de ésta, que englobe a la estudiada en la presente memoria. De manera más cercana a lo realizado en este trabajo, queda explicitar las tasas encontradas en el caso de los procesos de renovación, en términos de la distribución de espera F . Finalmente, también es interesante presentar algún tipo de evidencia numérica que indique posibles mejoras a lo realizado.

Por otra parte, también es importante notar que la mayoría de los resultados expuestos valen para cualquier coupling generado por (2.6). En esta dirección, es posible pensar en varias maneras de refinar las técnicas utilizadas; por ejemplo, dado que las funciones $\lambda(x) \vee \lambda(y)$ y $\lambda(x) + \lambda(y)$ actúan como intensidades *extremales* en cierto sentido, ¿es posible escoger un coupling que alterne entre ambas de manera óptima? Esto es, de qué forma puede escogerse un factor $\alpha \in [0, 1]$ (el que posiblemente depende del estado actual) de modo que $\beta(x, y) = \alpha(\lambda(x) \vee \lambda(y)) + (1 - \alpha)(\lambda(x) + \lambda(y))$ optimice la tasa a la cual se converge al equilibrio.

En resumen, este trabajo presenta una serie de resultados novedosos en el contexto de los procesos de Markov deterministas por pedazos, y abre una serie de interrogantes y nuevos caminos de investigación en esta área.

Bibliografía

- [1] S. Asmussen. *Applied Probability and Queues*. Applications of mathematics (Springer).: Stochastic modelling and applied probability. Springer, 2003. 1, 13, 17, 18, 24
- [2] R. Azaïs, J.-B. Bardet, A. Genadot, N. Krell, and P.-A. Zitt. Piecewise deterministic Markov process - recent results. ArXiv e-Print. 1
- [3] J.-B. Bardet, A. Christen, and J. Fontbona. Quantitative bounds for the renewal theorem with spread-out interarrivals. En preparación, 2013. 23
- [4] J.-B. Bardet, A. Christen, A. Guillin, F. Malrieu, and P.-A. Zitt. Total variation estimates for the tcp process. *ArXiv e-prints*, dec 2011. 1, 19, 29, 38
- [5] P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Statistics. Wiley, 2012. 3
- [6] D. Chafaï, F. Malrieu, and K. Paroux. On the long time behavior of the tcp window size process. *CoRR*, abs/0811.2180, 2008. 1, 13, 19
- [7] M. Davis. Piecewise-Deterministic Markov Process: A General Class of Non-Diffusion Stochastic Models. *Journal of the Royal Statistical Society. Series B (Methodology)*, 46(3):353–388, 1984. 1, 13
- [8] M. Davis. *Markov Models & Optimization*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 1993. 13
- [9] F. Dufour and O. Costa. Stability of piecewise-deterministic Markov processes. *SIAM Journal on Control and Optimization*, 37(5):1483–1502, 1999. 1
- [10] V. Dumas, F. Guillemin, and P. Robert. A markovian analysis of additive-increase multiplicative-decrease (aimd) algorithms. In *Advances in Applied Probability*, pages 85–111, 2002. 1, 18
- [11] S. Ethier and T. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. Wiley, 2009. 4, 5, 9
- [12] G. Folland. *Real analysis: modern techniques and their applications*. Pure and applied mathematics. Wiley, 1999. 6
- [13] F. Guillemin, P. Robert, and A. Zwart. AIMD algorithms and exponential functionals. Rapport de recherche RR-4447, INRIA, 2002. 18
- [14] M. Hairer. *Convergence of Markov Processes*, 2010.

- [15] M. Jacobsen. *Point Process Theory and Applications: Marked Point and Piecewise Deterministic Processes*. Probability and its applications. Springer, 2006. 13, 15
- [16] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate texts in mathematics. Springer, 1991. 3
- [17] T. Lindvall. *Lectures on the Coupling Method*. Dover Books on Mathematics. Dover Publications, 2012. 1
- [18] T. Ott, J. Kemperman, and M. Mathis. The stationary behavior of ideal TCP congestion avoidance. *progress, August, 1996*. 1, 13, 18
- [19] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Grundlehren der mathematischen Wissenschaften A series of comprehensive studies in mathematics. Springer, 1999. 3, 6, 7, 8, 9
- [20] G. O. Roberts and J. S. Rosenthal. Quantitative Bounds for Convergence Rates of Continuous Time Markov Processes. *Electronic Journal of Probability*, 1:1–21, May 1996. 1
- [21] H. Thórisson. *Coupling, stationarity and regeneration: with 27 illustrations*. Probability and Its Applications Series. Springer-Verlag, 2000. 1