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

UN CRITERIO DE UNICIDAD DE DISTRIBUCIONES CUASI-ESTACIONARIAS PARA UN PROCESO TRUNCADO DE NACIMIENTO Y MUERTE CON MUTACIONES

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL MATEMÁTICO

AMITAI SAMUEL LINKER GROISMAN

PROFESOR GUÍA:
JAIME SAN MARTIN

MIEMBROS DE LA COMISIÓN:
JAIME SAN MARTIN ARISTEGUI
SERVET MARTINEZ AGUILERA
DANIEL REMENIK ZISIS

SANTIAGO DE CHILE
ENERO 2014

RESUMEN DE LA MEMORIA
PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL MATEMÁTICO
POR: AMITAI LINKER
FECHA: 13/01/2014
PROF. GUÍA: JAIME SAN MARTÍN

UN CRITERIO DE UNICIDAD DE DISTRIBUCIONES CUASI-ESTACIONARIAS PARA UN PROCESO TRUNCADO DE NACIMIENTO Y MUERTE CON MUTACIONES

En la presente memoria se estudian las distribuciones cuasi-estacionarias (q.s.d.) de un proceso $\hat{Y}$ de nacimiento y muerte a tiempo continuo. Este proceso es absorbido al alcanzar una cantidad determinada de individuos, los cuales están caracterizados por rasgos fenotípicos, representados por elementos de un espacio métrico compacto. Cada uno de tales individuos puede morir o generar un nuevo individuo el cual puede poseer el mismo rasgo que su padre o mutar a uno de forma aleatoria. Las tasas del proceso pueden depender de la configuración de la población presente, la cual asumimos que se extingue casi-seguramente.

Para el estudio de las distribuciones cuasi estacionarias de $\hat{Y}$ se busca la existencia de una función acotada correspondiente a un vector propio por la derecha del semigrupo de transición $\{\hat{P}_t\}_{t\geq 0}$ del proceso, pues se prueba que en tal caso existe una única q.s.d., la cual es absolutamente continua respecto a una medida de referencia μ. La función antes mencionada es obtenida a partir del límite débil de vectores propios por la derecha para aproximaciones de $\{\hat{P}_t\}_{t\geq 0}$, bajo el único supuesto de que estas funciones son uniformemente acotadas. Además, para tales aproximaciones se prueba la existencia de vectores propios por la izquierda, los cuales corresponden a medidas de probabilidad que bajo el supuesto anterior convergen débilmente a la única distribución cuasi-estacionaria de $\hat{Y}$.

Se estudia finalmente un proceso Y correspondiente a la versión no absorbida de $\hat{Y}$, para el cual se prueban los mismos resultados bajo el supuesto adicional de que infinito se comporta como un estado de entrada.

“Las matemáticas pueden ser definidas como aquel tema en el cual no sabemos nunca lo que decimos ni si lo que decimos es verdadero”
- Bertrand Russell

Agradecimientos

Agradezco en primer lugar a mis padres, David y Jenny, por su cariño y apoyo incondicional. A ellos les debo todo lo que soy y las oportunidades que tengo para el futuro. También agradezco a mis hermanos Ioram y Yuval, mis partner de toda la vida, por ayudarme y soportarme durante todos estos años, y a quienes siempre puedo acudir para reír un rato. A mis primos, tíos y abuelos, por ayudarme a crecer y por creer en mi.

Agradezco a Natalia, mi gran compañera, por todo su amor y cariño, y por la increíble aventura que han sido estos 3 años juntos. Por apoyarme y soportar mis malos humores y ataques de histeria durante la realización de este trabajo, y por darme ánimos cada vez que fue necesario. Agradezco también a su familia por acogerme como si fuese uno de ellos.

Agradezco también a mi profesor guía, Jaime San Martín, por su ayuda fundamental en la realización de este trabajo y en mi desarrollo dentro del área de las probabilidades. Por guiarme tanto en lo académico como en lo personal, y por su tremenda simpatía y amistad. Del mismo modo agradezco a mi profesor co-guía, Servet Martínez, quien mantuvo siempre una excelente disposición para ayudarme. Agradezco también al profesor Daniel Remenik por aceptar amablemente el ser parte de la comisión de evaluación de este trabajo. Además agradezco a los profesores Andrew Hart del CMM, Loïc Hervé del INSA de Rennes, y Phill Pollett de la universidad de Queensland, por su increíble disposición.

A mis grandes y fundamentales amigos de la vida: Mota, Ilan y Ian, por compartir conmigo los buenos momentos y por el permanente apoyo y consejo que me han dado en los malos. Gracias por su amistad y cariño a los tres. A mis amigas Claudia y Alicia por estar siempre dispuestas a escucharme. A mis amigos del sombrero; Antonia, Cami, Cama, Sexy, Vicho, Meli, Seba, Cata, Petit (y Gabrielito), Julien, Fran, Amadeo, Pedrito, Java, Diego y Nati por los incontables momentos que hemos compartido. A mis amigos Yoav, Joel, Fede, Nicole, Vale, Domi, Danit, Johnny y Javi por su permanente apoyo y presencia en mi vida. También agradezco a mis buenos amigos durante estos años en la universidad, con quienes he tenido el honor de compartir largas jornadas de distracción: Conti, Pasti, Obando, Mauro, Tonho, Emilio, Vale, Tish, Víctor V., Nacho, Chino, Javi U., Marco, Chiri, Abelino, Zuñi, Angel y Funda. Especial mención tiene Nikolás Tapia, quien siempre ha estado ahí para ayudarme y a quien debo la comprensión de una parte importante de los conceptos involucrados en este trabajo.

Agradezco también a todos los funcionarios del DIM y del CMM, quienes han hecho de mi paso por el departamento algo mucho más sencillo, en especial a Gladys Cavallone, Eterin Jaña, Silvia Mariano, Oscar Mori, Luis Mella y María Inés Rivera.

Tabla de contenido

Capítulo 1

Introducción

Los procesos de Markov son utilizados frecuentemente como modelos dentro del área de las matemáticas aplicadas, siendo el teorema ergódico uno de los resultados más útiles para su estudio. Esta herramienta permite un análisis simple de las probabilidades de transición al largo plazo, para las cuales por lo general no se cuenta con expresiones explícitas, permitiendo así un estudio del comportamiento del sistema cuando éste ha alcanzado su "equilibrio". Sin embargo, los procesos para los cuales es posible aplicar el teorema corresponden únicamente a los procesos recurrentes, y por lo tanto nada se sabe a priori sobre el comportamiento al largo plazo de procesos transientes.

Una clase especial de procesos transientes son aquellos que dejan el espacio de estados casi seguramente en tiempo finito, evento al que nos referimos como *absorción*, y a los cuales llamamos procesos *absorbidos*. A pesar de no ser recurrentes, para muchos de tales procesos se observa un comportamiento estacionario previo a la absorción, como se ha descrito en modelos epidemiológicos (ver [24]). Este fenómeno es descrito mediante la existencia de equilibrios similares a las distribuciones estacionaras dentro de la teoría ergódica, los cuales en ciertos casos se comportan como *atractores*. Los equilibrios mencionados son conocidos como "distribuciones cuasi-estacionarias" (q.s.d.), y se definen como equilibrios para las probabilidades de transición condicionadas a la no-absorción del proceso, es decir, si X es un proceso de Markov absorbido y $\mathcal{T}$ es el tiempo de absorción de X, entonces π es una q.s.d. si

$$\mathbb{P}_\pi(X_t \in \cdot \mid \mathcal{T} > t) = \pi.$$

El objetivo de la presente memoria es el estudio de las propiedades de las q.s.d. para un proceso Y de nacimiento y muerte con mutaciones construido en [10], y para el cual la existencia de tales distribuciones fue probada en [4] bajo supuestos razonables. El proceso Y es introducido como una extensión de los procesos de nacimiento y muerte sobre $\mathbb{N}$ estudiados principalmente por Van Doorn, agregando al modelo la existencia de rasgos distintivos para los individuos de una población. De esta forma se otorga al modelo un nuevo grado de complejidad, ya que no sólo importa el tamaño de la población, sino que además es de interés saber la composición de la misma.

Desde el punto de vista matemático el estudio del proceso Y se presenta como un desafío considerable, pues su estructura es lo suficientemente complicada como para desestimar aproximaciones similares a las realizadas por Van Doorn en [30], y por otro lado, siendo un proceso de saltos sobre un espacio de estados continuo, no puede existir una medida de referencia respecto a la cual el semigrupo de transición sea absolutamente continuo, propiedad fundamental en buena parte de la literatura.

El enfoque utilizado en este trabajo es novedoso y corresponde a buscar la existencia de una función acotada que corresponda a un vector propio por la derecha del semigrupo de transición del proceso. Bajo la existencia de esta función, probamos que existe una única q.s.d. la cual es absolutamente continua respecto a una medida de referencia μ, y aún más, si π es la q.s.d. encontrada, y η es un estado inicial cualquiera, entonces

$$\mathbb{P}_\eta(Y_t \in \cdot \mid \mathcal{T} > t) \xrightarrow[t\to\infty]{} \pi.$$

La función antes mencionadas será obtenida a partir del límite débil de una sucesión de funciones que corresponden a vectores propios por la derecha para procesos aproximados, bajo el supuesto de que tales funciones son uniformemente acotadas. Además, para tales procesos encontraremos distribuciones cuasi-estacionarias, las cuales bajo el supuesto anterior convergerán a la única q.s.d. de Y.

1.1. Organización del trabajo

En el segundo capítulo de esta memoria construiremos el proceso Y a partir de una cadena de Markov $\{(\bar{Y}_n, \tau_n)\}_{n\in\mathbb{N}}$, en la cual τ_n representa el tiempo del n-ésimo salto, y $\bar{Y}_n$ representa el valor de Y en ese instante. A continuación introducimos el kernel de transición Q y el generador infinitesimal G del proceso y mostramos que bajo una serie de hipótesis adecuadas (utilizadas originalmente en [4]), el proceso no posee explosiones y es absorbido exponencialmente.

En el tercer capítulo introducimos algunos conceptos elementales sobre distribuciones cuasi-estacionarias y procesos absorbidos, siendo de especial interés el Teorema 3.3 según el cual las distribuciones cuasi-estacionarias corresponden a vectores propios por la izquierda para el semigrupo de transición. Cerraremos este capítulo entregando algunos resultados básicos de existencia de q.s.d. para procesos con la propiedad de Feller, y su aplicación al proceso Y.

En el cuarto capítulo introducimos el concepto de φ-irreducibilidad, junto con una medida μ cuya construcción es explícita y que resulta ser una medida maximal de irreducibilidad. A continuación introducimos el concepto de medidas *bien concentradas*, las cuales incluyen a μ y a las q.s.d. y que son de interés pues para tales medidas los elementos del semigrupo de transición están bien definido como operadores de L^p en sí mismo. Las medidas bien concentradas permiten enunciar el Teorema 4.2, que es de carácter fundamental en este trabajo, pues enuncia la continuidad de tales operadores. Terminaremos este capítulo enunciando los principales resultados de Λ^*-clasificación, los cuales son una extensión de los resultados de

Tweedie para procesos a tiempo discreto, y a partir de los cuales deduciremos las consecuencias de la existencia de un vector por la derecha acotado para el semigrupo de transición.

El quinto capítulo está dedicado al estudio del proceso truncado $\hat{Y}$, que corresponde a absorber el proceso una vez que se ha alcanzado una población de cierto tamaño, y para el cual veremos que se mantienen las propiedades de Y obtenidas en los capítulos anteriores. Se definirá para $\hat{Y}$ una partición del espacio de estados que permite la construcción de una secuencia de procesos $\hat{W}^n$ cuyas tasas toman valores constantes en cada elemento de la partición y que "aproximan" a $\hat{Y}$.

Veremos que estos procesos se comportan como si el espacio de estados fuese discreto, y por lo tanto deduciremos mediante el Teorema 5.1 (que es una adaptación del Teorema 1 en [21]), la existencia y unicidad de distribuciones cuasi-estacionarias π_n para ellos. Construiremos además para cada n una función f_n acotada que corresponde a un vector propio por la derecha para el semigrupo de transición de $\hat{W}^n$.

Probaremos a continuación el Teorema 5.2, el cual indica que si la secuencia $\{f_n\}_{n\in\mathbb{N}}$ es uniformemente acotada, entonces ésta converge a la función f buscada, y que además la secuencia $\{\pi_n\}_{n\in\mathbb{N}}$ converge a la única q.s.d. π de $\hat{Y}$. Concluiremos el capítulo presentando algunos resultados numéricos y una extensión del teorema para el caso en que las tasas de transición no son continuas.

En el capítulo 6 se introducen algunos resultados básicos para procesos de nacimiento y muerte en $\mathbb{N}$, obtenidos por Van Doorn en [30], y luego se realiza para Y la misma construcción de procesos W^n bajo el supuesto de que Y *vuelve rápido desde infinito*. Como consecuencia se deduce un resultado similar al Teorema 5.2 para este caso.

Capítulo 2

El proceso Y

En el presente capítulo introducimos el proceso de nacimiento y muerte Y construido en [4], el cual toma valores en un espacio de estados $\mathcal{A}$ cuya estructura se estudia en detalle. Siendo Y un proceso de saltos, estudiaremos su comportamiento a través de sus tasas de transición, las cuales asumiremos conocidas y que cumplen un conjunto determinado de hipótesis, el cual notaremos como (H).

El proceso será construido a partir de una cadena de Markov, la cual definiremos trayectorialmente, y utilizando tal definición mostraremos que es la única solución de una ecuación integral específica. Aunque la construcción de Y resulta algo engorrosa, ésta permite acotar trayectorialmente el proceso y probar con suficiente formalidad que un cierto operador G corresponde a su generador débil.

2.1. El espacio de estados

A diferencia de los modelos de nacimiento y muerte estudiados normalmente, donde el espacio de estados corresponde a la cantidad de individuos observados, en el presente modelo, este espacio debe ser capaz de almacenar de forma sencilla el tamaño de la población y el detalle de los rasgos presentes en ella en cada momento. Para ello introduciremos en esta sección el espacio $\mathcal{A}$ construido en [4].

Sea $(\mathbb{T}, d_{\mathbb{T}}, \mathcal{B}_{\mathbb{T}})$ en que $\mathbb{T}$ es un conjunto, $d_{\mathbb{T}}$ es una métrica en $\mathbb{T}$ bajo la cual es compacto y de diámetro 1, y $\mathcal{B}_{\mathbb{T}}$ es la σ-álgebra de los Borelianios para la topología generada por $d_{\mathbb{T}}$. Consideramos $\mathbb{T}$ como el espacio de rasgos, cuya metrizabilidad y compacidad son restricciones naturales.

Observación Gracias a su compacidad $\mathbb{T}$ es separable, posee una base numerable de abiertos, $\mathcal{B}_{\mathbb{T}}$ es numerablemente generada, y podemos construir una relación de orden total $\leq$ en $\mathbb{T}$ que sea $\mathcal{B}_{\mathbb{T}}$-medible (ver [4, Sección 2]).

Definición 2.1 *Sea $\mathcal{M}(\mathcal{B}_{\mathbb{T}})$ el espacio de medidas sobre $\mathcal{B}_{\mathbb{T}}$, definimos el conjunto de estados del proceso, $\mathcal{A} \subseteq \mathcal{M}(\mathcal{B}_{\mathbb{T}})$ como*

$$\mathcal{A} := \left\{ \sum_{i=1}^{n} \delta_{x_i} \;\; n \in \mathbb{N},\; x_1, x_2, \ldots, x_n \in \mathbb{T} \right\} \cup \{0\},$$

donde 0 es la medida nula y cada δ_{x_i} es la delta de Dirac centrada en x_i. Llamaremos configuraciones a los elementos del espacio $\mathcal{A}$ e individuos a los elementos de la forma δ_x (cuyo rasgo asociado es x).

Los elementos en $\mathcal{A}$ permiten una interpretación sencilla de la composición de la población, por ejemplo un elemento de la forma $2\delta_x + 3\delta_y + \delta_z$ indica que la población presenta seis individuos, de los cuales dos presentan el rasgo x, tres el rasgo y y uno el rasgo z.

Aunque los elementos en $\mathcal{A}$ resumen la totalidad de la información necesaria respecto a las configuraciones, en ocasiones será necesario hacer referencia a aspectos específicos de las mismas, tales como los rasgos presentes y su frecuencia, cuántos individuos presenta un cierto rasgo, etc. Para manejar estos aspectos se utilizarán las siguientes definiciones, en las cuales es importante notar que dada una configuración η y un conjunto $B \subseteq \mathbb{T}$, $\eta(B)$ indica la cantidad de individuos en η que presentan un rasgo en B.

Definición 2.2 *Para cada configuración $\eta \in \mathcal{A}$ no vacía y para cada $y \in \mathbb{T}$ y $k \in \mathbb{N}$ definimos*

1. $\eta_y := \eta(\{y\}) \in \mathbb{N}$, *la cantidad de individuos que presentan el rasgo y;*
2. $\{\eta\} := \{y \in \mathbb{T},\ \eta_y > 0\} \subseteq \mathbb{T}$, *el conjunto de rasgos presentes en η;*
3. $\#\eta := Card(\{\eta\}) \in \mathbb{N}$, *la cantidad de rasgos presentes en η;*
4. $y_{\eta,1} := \min\{y \in \{\eta\}\} \in \mathbb{T}$, *el menor rasgo presente en η;*
5. $y_{\eta,k} := \min\{y \in \{\eta\},\ y_{\eta,k-1} < y\} \in \mathbb{T}$, *el k-ésimo ($k \leq \#\eta$) rasgo más pequeño en η;*
6. $\vec{\eta} := (y_{\eta,1}, y_{\eta,2}, \ldots, y_{\eta,\#\eta}) \in \mathbb{T}^{\#\eta}$, *el conjunto ordenado de rasgos presentes en η;*
7. $\bar{\eta} := (\eta_{y_{\eta,1}}, \eta_{y_{\eta,2}}, \ldots, \eta_{y_{\eta,\#\eta}}) \in \mathbb{N}^{\#\eta}$, *el vector de frecuencias de los rasgos ordenados;*
8. $S_k(\eta) := \sum_{i=1}^{k} \bar{\eta}_i \in \mathbb{N}$, *la frecuencia acumulada hasta el rasgo k.*

De la definición anterior serán de particular utilidad los elementos $\eta_y, \{\eta\}, \vec{\eta}$ y $\bar{\eta}$, pues permiten manejar conceptos fundamentales de una configuración. Notemos que ninguna de las definiciones anteriores da cuenta de la cantidad total de individuos, la cual manejaremos mediante la siguiente función.

Definición 2.3 *Definimos la función $\|\cdot\| : \mathcal{A} \to \mathbb{N}$ como $\|\eta\| := \int 1 d\eta = \eta(\mathbb{T})$, la cantidad de individuos de la configuración η. A partir de esta función se definen*

- $\mathcal{A}_m := \|\cdot\|^{-1}(\{m\})$, *las configuraciones con m individuos;*
- $\mathcal{A}_{\leq m} := \mathcal{A}_1 \cup \mathcal{A}_2 \cup \cdots \cup \mathcal{A}_m$, *las configuraciones no vacías con m individuos o menos;*
- $\mathcal{A}_{\geq m} := \mathcal{A}_m \cup \mathcal{A}_{m+1} \cup \cdots$, *las configuraciones con m individuos o más;*
- $\mathcal{A}^{-0} := \mathcal{A}_{\geq 1}$, *el conjunto de configuraciones no vacías.*

Utilizando la función anterior podemos introducir una última expresión $\vec{H}(\eta)$, la cual inyecta $\mathcal{A}_m$ en $\mathbb{T}^m$.

Definición 2.4 *Sean $\eta \in \mathcal{A}$ y $k \in \mathbb{N}$. Se define el vector $\vec{H}(\eta) \in \mathbb{T}^{\|\eta\|}$ en que cada componente $\vec{H}_k(\eta)$ es de la forma*

$$\vec{H}_k(\eta) \;=\; y_j \quad con \quad j \leq \#\eta \quad y \quad k \in (S_{j-1}(\eta), S_j(\eta)].$$

Es directo verificar que $\vec{H}(\eta)$ corresponde al vector de los $\|\eta\|$ rasgos presentes en η, ordenados y repetidos según su frecuencia. A modo de ejemplo, y con el fin de ilustrar las definiciones anteriores consideremos $\eta = 2\delta_x + 3\delta_y + \delta_z$ en que $x \leq y \leq z$. Para esta configuración se tiene

$$\{\eta\} = \{x, y, z\}, \quad \vec{\eta} = (x, y, z), \quad \bar{\eta} = (2, 3, 1), \quad \|\eta\| = 6 \;\text{ y }\; \vec{H}(\eta) = (x, x, y, y, y, z).$$

Ya habíamos comentado que para cada $m \in \mathbb{N}$, $\vec{H}$ inyecta de forma natural $\mathcal{A}_m$ en $\mathbb{T}^m$, y veremos a continuación que dotando $\mathcal{A}_m$ de la topología adecuada, estos espacios poseen propiedades similares. Para ello notemos que para cada $n \in \mathbb{N}$, $\mathcal{A}_n \subseteq n\mathcal{P}(\mathcal{B}_\mathbb{T})$ donde $\mathcal{P}(\mathcal{B}_\mathbb{T})$ es el conjunto de medidas de probabilidad en $\mathcal{B}_\mathbb{T}$. Podemos dotar entonces a $\frac{1}{n}\mathcal{A}_n$ de la métrica de Lévy-Prokhorov, d_P, cuya topología coincide en este caso con la topología débil-* puesto que $\mathbb{T}$ es separable (ver [1, Teorema 6.8]).

Proposición 2.1 *Cada conjunto $\frac{1}{m}\mathcal{A}_m$ dotado de la métrica de Lévy-Prokhorov es compacto.*

Demostración. Sea $\{\eta_k\}_{k\in\mathbb{N}}$ una sucesión en $\frac{1}{m}\mathcal{A}_m$. Por definición la secuencia $\{\vec{H}(m\eta_k)\}_{k\in\mathbb{N}}$ se encuentra contenida en $\mathbb{T}^m$ que es compacto, y por lo tanto posee una subsucesión convergente a algún elemento $(x_1, x_2, \ldots, x_m)$. Para mostrar que $\eta_k \to \eta$ salvo subsucesión, en que $\eta = \frac{1}{m}\sum_{i=1}^m \delta_{x_i}$, notemos que para toda $f : \mathbb{T} \to \mathbb{R}$ continua se tiene

$$\int_\mathbb{T} f d\eta_k \;=\; \frac{1}{m}\sum_{i=1}^m f\left(\vec{H}_i(m\eta_k)\right) \;\xrightarrow[k\to\infty]{}\; \frac{1}{m}\sum_{i=1}^m f(x_i) \;=\; \int_\mathbb{T} f d\eta,$$

en que hemos utilizado la continuidad de f y la convergencia de cada x_i^k a x_i para concluir la convergencia de la suma. Así, por definición de convergencia débil-*, $\{\eta_k\}_{k\in\mathbb{N}}$ posee subsucesión convergente.

$\square$

Dado que el conjunto $\mathcal{A}$ es la unión numerable y disjunta de los conjuntos $\mathcal{A}_m$ y $\{0\}$ podemos extender d_P a una métrica d sobre $\mathcal{A}$ cuya topología es equivalente a la de la convergencia débil-*. En efecto, basta tomar

$$d(\eta_1, \eta_2) \;:=\; \begin{cases} d_P\left(\frac{\eta_1}{\|\eta_1\|}, \frac{\eta_2}{\|\eta_2\|}\right) & \text{si} \quad \|\eta_1\| = \|\eta_2\| \\ 3\big|\|\eta_1\| - \|\eta_2\|\big| & \text{si} \quad \|\eta_1\| \neq \|\eta_2\| \end{cases},$$

bajo la cual $\mathcal{A}$ es localmente compacto y cada $\mathcal{A}_m$ es un abierto y un cerrado. Si dotamos además a este conjunto de la σ-álgebra de los borelianos, $\mathcal{B}$, es claro que los elementos introducidos en las Definiciones 2.2, 2.3 y 2.4 son medibles.

Observación $\mathcal{A}$ es completo. En efecto, dada una sucesión $\{\eta_n\}_{n\in\mathbb{N}}$ de Cauchy, existe $N \in \mathbb{N}$ tal que $\forall n, m \geq N, d(\eta_n, \eta_m) \leq \frac{1}{2}$, luego por definición $\{\eta_m\}_{m\geq N} \subseteq \mathcal{A}_k$ para algún k, y como este conjunto es compacto, la sucesión converge.

2.2. Las tasas de transición

Ya introducido el conjunto $\mathcal{A}$ de estados, definiremos las tasas de transición del proceso siguiendo [4], en que utilizamos la notación $\mathcal{C}(B)$ para referirnos al conjunto de funciones continuas sobre un conjunto B a valores en $\mathbb{R}$.

Definición 2.5 *Sean $b, m,$ y λ tres elementos en $\mathcal{C}(\mathbb{T} \times \mathcal{A})$, a los cuales nos referiremos como las tasas de transición. Estas tres funciones toman valores estrictamente positivos en el conjunto $\mathbb{T} \times \mathcal{A}^{-0}$ y valores nulos fuera de él. Definimos además una medida de probabilidad σ sobre $(\mathcal{B}_{\mathbb{T}})$, la cual no posee átomos y cuyo soporte es $\mathbb{T}$, y la densidad $g \in \mathcal{C}(\mathbb{T}^2)$ a valores estrictamente positivos y tal que*

$$\forall y \in \mathbb{T}, \qquad \int_{\mathbb{T}} g(y, z) d\sigma(z) = 1.$$

El proceso que se busca modelar corresponde a un proceso de saltos a valores en $\mathcal{A}$, en que los saltos corresponden a eventos en los que cambia la población. Tales eventos pueden ser el nacimiento de un nuevo individuo (ya sea con el mismo rasgo del padre o no), o la muerte del mismo. Cada individuo aporta una tasa para el acontecimiento de alguno de tales eventos de la siguiente manera.

- Cada individuo δ_y en η aporta una tasa $b(y, \eta)$ para el nacimiento de un individuo con su propio rasgo (clonación);
- cada individuo δ_y en η aporta una tasa $\lambda(y, \eta)$ para su propia muerte;
- cada individuo δ_y en η aporta una tasa $m(y, \eta)$ para el nacimiento de un individuo con un rasgo distinto al suyo (mutación).

Si a partir de δ_y se produce un nacimiento con mutación, el rasgo del nuevo individuo se elige con distribución de probabilidad dada por la medida $g(y, \cdot)d\sigma$, luego podemos interpretar g como la distribución de mutaciones respecto a una medida σ que representa la distribución natural de los rasgos en $\mathbb{T}$.

Para $y \in \mathbb{T}$ y $\eta \in \mathcal{A}$ utilizamos la notación $b_y(\eta)$ en lugar de $b(y, \eta)$. De igual forma utilizamos $m_y(\eta)$ y $\lambda_y(\eta)$ en lugar de $m(\eta, y)$ y $\lambda(\eta, y)$. El aporte de las tasas de cada individuo para la ocurrencia de algún evento se ve resumido en el kernel de transición Q que definimos a continuación.

Definición 2.6 *Se define el kernel de transición $Q : \mathcal{A} \times \mathcal{B}_{\mathcal{A}} \to \mathbb{R}_+$ como*

$$Q(\eta, B) \ := \sum_{\substack{y \in \{\eta\} \\ \eta + \delta_y \in B}} \eta_y b_y(\eta) \ + \sum_{\substack{y \in \{\eta\} \\ \eta + \delta_y \in B}} \eta_y \lambda_y(\eta) \ + \sum_{y \in \{\eta\}} \eta_y m_y(\eta) \int_{\eta + \delta_z \in B} g(y, z) d\sigma(z).$$

Será de particular interés el caso en que $B = \mathcal{A}$, en cuyo caso utilizamos la notación

$$Q(\eta) := Q(\eta, \mathcal{A}) = \sum_{y \in \{\eta\}} \eta_y [b_y(\eta) + m_y(\eta) + \lambda_y(\eta)].$$

Observación Para cada $\eta \in \mathcal{A}$, $Q(\eta, \cdot)$ es una medida, y gracias a su construcción utilizando funciones medibles, para cada $B \in \mathcal{B}_{\mathcal{A}}$, $Q(\cdot, B)$ es $\mathcal{B}_{\mathcal{A}}$-medible, luego Q es efectivamente un kernel. Además, gracias a la definición 2.5 para cada $\eta \neq 0$ se tiene $Q(\eta) > 0$ y $Q(0) = 0$.

Debido a la construcción de $Q(\cdot, \mathcal{A})$ utilizando funciones continuas, es de esperar que tal función también lo sea. Aplicando el siguiente resultado, probado en [4], a la función $b + m + \lambda$ muestra que esto efectivamente ocurre.

Proposición 2.2 *[4, Lema 2.6] Sea $F : \mathcal{A} \times \mathbb{T} \to \mathbb{R}$ una función continua, entonces la función $\tilde{F} : \mathcal{A} \to \mathbb{R}$ es continua, donde*

$$\tilde{F}(\eta) \ := \sum_{y \in \{\eta\}} F(\eta, y) \eta_y.$$

2.3. Construcción de Y

En lo que resta del capítulo consideraremos los espacios de medida $(\mathbb{R}_+, \mathcal{B}_{\mathbb{R}}, \mu_{\mathbb{R}})$, $(\mathbb{N}, \mathcal{B}_{\mathbb{N}}, \mu_{\mathbb{N}})$ y $(\mathbb{T}, \mathcal{B}_{\mathbb{T}}, \sigma)$, en que $\mathcal{B}_{\mathbb{R}}$ es la tribu Boreliana, $\mu_{\mathbb{R}}$ es la medida de Lebesgue, $\mathcal{B}_{\mathbb{N}}$ es la σ-álgebra discreta en $\mathbb{N}$, $\mu_{\mathbb{N}}$ es la medida cuentapuntos en $\mathbb{N}$, y $(\mathbb{T}, \mathcal{B}_{\mathbb{T}}, \sigma)$ es como en la Definición 2.5. Introduciremos a continuación un conjunto de medidas de Poisson sobre un espacio de intensidades E a partir de las cuales construiremos Y.

Definición 2.7 *A partir de los espacios anteriores se definen*

- *El espacio de intensidades $E := \mathbb{R}_+ \times \mathbb{N} \times \mathbb{T} \times \mathbb{R}_+$, dotado de la σ-álgebra producto $\Sigma := \mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{N}} \otimes \mathcal{B}_{\mathbb{T}} \otimes \mathcal{B}_{\mathbb{R}}$ y la medida producto $\mu_E := \mu_{\mathbb{R}} \otimes \mu_{\mathbb{N}} \otimes \sigma \otimes \mu_{\mathbb{R}}$;*

- *el espacio de probabilidad del proceso $(\Omega, \mathcal{F}, \mathbb{P})$;*

- *las medidas de Poisson $M_c, M_m, M_d : \Omega \times \Sigma \to \mathbb{N} \cup \{0\}$ independientes de a pares con intensidad μ_E.*

Recordemos (ver [9, Definición 3.2]) que $N : \Omega \times \Sigma \to \mathbb{N}$ es una medida de Poisson con intensidad μ_E si se cumple,

1. Para cada $A \in \Sigma$, $N(\cdot, A)$ es una variable aleatoria de Poisson con parámetro $\mu_E(A)$;

2. para cada $\omega \in \Omega$, $N(\omega, \cdot)$ es una medida puntual;
3. si $A_1, A_2 \in \Sigma$ cumplen $A_1 \cap A_2 = \emptyset$ entonces $N(\cdot, A_1)$ y $N(\cdot, A_2)$ son independientes.

Dentro del espacio de intensidades reflejamos las tasas definidas en 2,5 a partir de la construcción de las siguientes *ventanas*.

Definición 2.8 *Sea $\eta \in \mathcal{A}$ cualquiera, a partir de esta configuración se definen los siguientes conjuntos,*

- $E_c(\eta) := \big\{(j, z, \theta) \in \mathbb{N} \times \mathbb{T} \times \mathbb{R}_+ \mid\ 1 \leq j \leq \|\eta\|,\ 0 \leq \theta \leq b_{H^j(\eta)}(\eta)\big\}$;
- $E_m(\eta) := \big\{(j, z, \theta) \in \mathbb{N} \times \mathbb{T} \times \mathbb{R}_+ \mid 1 \leq j \leq \|\eta\|,\ 0 \leq \theta \leq m_{H^j(\eta)}(\eta) g(y, z)\big\}$;
- $E_d(\eta) := \big\{(j, z, \theta) \in \mathbb{N} \times \mathbb{T} \times \mathbb{R}_+ \mid\ 1 \leq j \leq \|\eta\|,\ 0 \leq \theta \leq \lambda_{H^j(\eta)}(\eta)\big\}$.

Un cálculo sencillo muestra que $\mu_{\mathbb{N}} \otimes \sigma \otimes \mu_{\mathbb{R}} \left[E_c(\eta) \cup E_m(\eta) \cup E_d(\eta))\right] = Q(\eta)$, luego estos conjuntos efectivamente reflejan la tasa con que ocurrirá el siguiente evento. Utilizando estas definiciones construimos una cadena de Markov $\{(\bar{Y}_n, \tau_n)\}_{n \in \mathbb{N}}$ de la siguiente forma.

Definición 2.9 *Sean $Y_0 : \Omega \to \mathcal{A}$ un elemento aleatorio independiente de M_c, M_m y M_d introducidas como en la definición 2.7. Se construye la sucesión $\{(\bar{Y}_n, \tau_n)\}_{n \in \mathbb{N}} \subseteq \mathcal{A} \times \mathbb{R}$ de forma recursiva como*

$$
\begin{array}{lll}
- \ \tau_0 & := & 0 \\
- \ \bar{Y}_0 & := & Y_0 \\
- \ E_x^{k,t} & := & (\tau_k, t] \times E_x(\bar{Y}_k) \qquad \textit{con } x = c, m, d \\
- \ \tau_{k+1} & := & \inf\{t > \tau_k,\ M_c(E_c^{k,t}) + M_m(E_m^{k,t}) + M_d(E_d^{k,t}) > 0\} \\
- \ E_x^k & := & (\tau_k, \tau_{k+1}] \times E_x(\bar{Y}_k) \qquad \textit{con } x = c, m, d \\
- \ v_x^k & := & \int_{E_x^k} \delta_{H^j(\bar{Y}^k)} M_x(dt, dj, dz, d\theta) \qquad \textit{con } x = c, d \\
- \ v_m^k & := & \int_{E_x^k} \delta_z M_m(dt, dj, dz, d\theta) \\
- \ \bar{Y}_{k+1} & := & \bar{Y}_k + v_c^k + v_m^k - v_d^k
\end{array}
$$

La mayor parte de los elementos introducidos en la definición anterior tienen como única finalidad abreviar la notación de la construcción de la cadena, cuya interpretación es como sigue;

- supongamos que $\bar{Y}_k = \eta$ y que $\tau_k = s$. Se fijan tres *fibras* en E, las cuales corresponden a $E_m(\eta), E_c(\eta)$ y $E_d(\eta)$;
- una vez fijadas las fibras anteriores se generan las ventanas $E_x(\eta) \times (s, t]$, cada una asociada a su respectiva medida M_x. Se aumenta el valor de t hasta que en alguna de ellas, la medida es no nula (esto es, marque un evento);
- el valor de t para el cual se marca un evento corresponde a τ_{k+1}. A continuación se generan los vectores v_x cuyo propósito es mostrar qué individuo se agrega o muere a $\bar{Y}_k$ para generar $\bar{Y}_{k+1}$.

En general nos interesarán únicamente las variables $\{\tau_n\}_{n\in\mathbb{N}}$ y $\{\bar{Y}_n\}_{n\in\mathbb{N}}$, las cuales forman una cadena de Markov como se indica en el siguiente resultado, cuya demostración es de carácter más bien técnico y se encuentra en el Anexo.

Proposición 2.3 *1. La sucesión $\{(\bar{Y}_n, \tau_n)\}_{n\in\mathbb{N}}$ es cadena de Markov homogénea;*

2. las variables $\Delta\tau_{k+1}\big|_{\bar{Y}_k}$ en que $\Delta\tau_{k+1} = \tau_{k+1} - \tau_k$, son independientes de a pares y tienen distribución exponencial con parámetro $Q(\bar{Y}_k)$;

3. los elementos condicionados $\bar{Y}_{n+1}\big|_{\bar{Y}_n}$ tienen como distribución $\frac{Q(\bar{Y}_n,\cdot)}{Q(\bar{Y}_n)}$;

4. para todo $n \in \mathbb{N}$ las variables $\bar{Y}_{n+1}\big|_{\bar{Y}_n}$ y $\Delta\tau_{n+1}\big|_{\bar{Y}_n}$ son independientes;

5. la sucesión $\{\bar{Y}_n\}_{n\in\mathbb{N}}$ es cadena de Markov homogénea.

De esta forma, la cadena $\{\bar{Y}_n\}_{n\in\mathbb{N}}$ puede interpretarse como el proceso Y que se busca construir mirado en los tiempos de salto. En efecto, suponiendo conocido $\bar{Y}_n = \eta$, $\delta\tau_{n+1}$ corresponde a un tiempo de espera exponencial de parámetro $Q(\eta)$ y que es independiente de $\bar{Y}_{n+1}$, el cual tiene distribución $\frac{Q(\eta,\cdot)}{Q(\eta)}$. Notemos que las variables τ_n, correspondientes a los tiempos de salto, permiten introducir naturalmente la siguiente familia de variables, que interpretamos como la cantidad de saltos ocurridos en un intervalo de tiempo.

Definición 2.10 *Sean $\{\tau_n\}_{n\in\mathbb{N}}$ como en la Definición 2.9. Se define la familia de variables aleatorias $\{N_t\}_{t\geq 0}$ como*

$$N_t := \sup\{n \in \mathbb{N} \cup \{0\},\ \tau_n \leq t\}.$$

Utilizando esta familia de variables aleatorias, podemos definir el proceso Y_t como sigue:

Definición 2.11 *Se define la familia $\{Y_t\}_{t\geq 0}$ de elementos aleatorios en $\mathcal{A}$ como*

$$Y_t \ := \ \bar{Y}_{N_t},$$

y en que $\bar{Y}_\infty := 0$ donde 0 es la configuración vacía.

La siguiente es una consecuencia directa de la Proposición 2.3 y de [11, Teorema 3.8]:

Proposición 2.4 *El proceso $\{Y_t\}_{t\geq 0}$ es un proceso de Markov homogéneo cuyos tiempos de salto corresponden a las variables $\{\tau_n\}_{n\in\mathbb{N}}$ y en que $Y_{\tau_n} = \bar{Y}_n$.*

2.4. Hipótesis sobre las tasas

Con el fin de lograr un cierto nivel de control sobre el comportamiento de Y asumiremos que sus tasas de transición cumplen las siguientes hipótesis, de las cuales se desprenden propiedades importantes, tales como la absorción casi segura a la configuración nula, la ausencia de explosiones y la expresión del proceso como única solución de una ecuación integral.

Hipótesis 1 (**H1**). *Con el fin de evitar un crecimiento descontrolado de la población, asumimos*

$$\sup_{y\in\mathbb{T}}\ \sup_{\eta\in\mathcal{A}}\ (b_y(\eta)+m_y(\eta))\ <\ \infty.$$

Hipótesis 2 (**H2**). *Para controlar las propiedades de martingala, asumiremos que existen $p \geq 1$ y $c > 0$ tales que $\forall \eta \in \mathcal{A}$*

$$\sup_{y\in\mathbb{T}} \lambda_y(\eta)\ \leq\ c\|\eta\|^p.$$

Definición 2.12 *Introducimos las cotas al largo plazo para las tasas de Y como*

$$\begin{aligned}\lambda_* &:= \liminf_{k\to\infty}\ \inf_{\|\eta\|=k}\ \inf_{y\in\{\eta\}} \lambda_y(\eta);\\ \Gamma^* &:= \limsup_{k\to\infty}\ \sup_{\|\eta\|=k}\ \sup_{y\in\{\eta\}} (b_y(\eta)+m_y(\eta)).\end{aligned}$$

Hipótesis 3 (**H3**) *Con el fin de asegurar la absorción del proceso a la configuración nula asumiremos que*

$$\Gamma^*\ <\ \lambda_*.$$

Nos referimos al conjunto de hipótesis (H1), (H2) y (H3) sencillamente como (H).

2.4.1. Explosiones del proceso

Como una primera aplicación de las hipótesis (H) mostraremos que el proceso Y_t no posee explosiones, es decir, $N_t < \infty$ casi-seguramente. Para ello seguiremos los resultados demostrados en [4], adaptándolos a la construcción de Y realizada en este trabajo. Definimos entonces la sucesión de variables aleatorias $\{(\bar{Z}_n, \bar{\tau}_n)\}_{n\in\mathbb{N}}$ de forma similar a lo hecho para $\{(\bar{Y}_n, \tau_n)\}_{n\in\mathbb{N}}$ como sigue.

Definición 2.13 *Sea $k \in \mathbb{N}$ cualquiera, definimos el conjunto $E_z(k)$ como*

$$E_z(k)\ :=\ \{1,2,\ldots,k\}\times\mathbb{T}\times[0,C^*],$$

en que hemos tomado C^* de la forma

$$C^*\ :=\ \left(\sup_{\eta\in\mathcal{A}}\sup_{y\in\{\eta\}}\ (b_y(\eta)+m_y(\eta))\right)\left(\sup_{y,z\in\mathbb{T}} g(y,z)\right),$$

y que es finito gracias a la hipótesis (H1) y la continuidad de g sobre el compacto $\mathbb{T}^2$. Definimos a continuación la sucesión $\{(\bar{Z}_n, \bar{\tau}_n)\}_{n\in\mathbb{N}}$ como

- $\bar{Z}_0 \;=\; \|Y_0\|$;
- $\bar{\tau}_0 \;=\; 0$;
- $\bar{\tau}_{n+1} \;=\; \inf\{t > \bar{\tau}_n,\ M_c((\bar{\tau}_n, t]\times E_z(\bar{Z}_n)) + M_m((\bar{\tau}_n, t]\times E_z(\bar{Z}_n)) > 0\}$;
- $\bar{Z}_{n+1} \;=\; \bar{Z}_n + 1$.

De forma similar al caso de $\{(\bar{Y}_n, \bar{T}_n)\}_{n\in\mathbb{N}}$ se tiene que $\{(\bar{Z}_n, \bar{\tau}_n)\}_{n\in\mathbb{N}}$ es una cadena de Markov en que las variables $\Delta\bar{\tau}_{n+1}|_{\bar{Z}_n}$ son independientes y con distribución exponencial de parámetro $2\bar{Z}_n C^*$. Definiremos entonces el proceso Z de forma análoga a Y.

Definición 2.14 *Dada la sucesión $\{(\bar{Z}_n, \bar{\tau}_n)\}_{n\in\mathbb{N}}$ y $t > 0$ se define la familia de variables aleatorias N_t^Z como*

$$N_t^Z := \sup\{n \in \mathbb{N} \cup \{0\},\ \bar{\tau}_n \leq t\},$$

a partir de las cuales se define

$$Z_t \ := \ \bar{Z}_{N_t^Z},$$

y en que $\bar{Z}_\infty = \infty$.

Notemos que este es un proceso de nacimiento puro en $\mathbb{N}$ que acota el tamaño de la población en Y, como veremos en el próximo resultado mencionado en [4].

Proposición 2.5 *Para todo $t \geq 0$ se cumple $\mathbb{P}$-casi seguramente*

$$\|Y_t\| \ \leq \ Z_t\,.$$

Demostración. Sea $m \in \mathbb{N}$ cualquiera. Utilizaremos la notación $\mathbb{P}_m$ para referirnos a la medida $\mathbb{P}$ condicionada al evento $\|Y_0\| = m$, bajo el cual $\bar{Z}_n = m + n$ y cada variable $\Delta\bar{\tau}_{n+1}$ tiene distribución exponencial de parámetro $2(m+n)C^*$. Como podemos ver en [18, Corolario 1], el crecimiento de tales parámetros es suficientemente lento como para asegurar que $\mathbb{P}$-c.s. el proceso Z_t no presente explosiones.

Consideremos ahora una realización de ambos procesos en que Z_t no posee explosiones, y sea $(t_1, t_2, \ldots)$ una enumeración ordenada de los tiempos de salto de los procesos Y_t y Z_t hasta el tiempo s. Por definición, tanto Y_t como Z_t son constantes en los intervalos de la forma $[t_i, t_{i+1})$, luego basta mostrar por inducción que para cada $k \geq 0$ se cumple $\|Y_{t_k}\| \leq Z_{t_k}$.

El caso base se tiene trivialmente pues $t_0 = 0$ y por definición $\|Y_0\| = Z_0$. Para mostrar el caso $k+1$ sabiendo que se cumple el caso k, notemos que hay dos eventos posibles,

1. $\|Y_{t_{k+1}}\| \leq \|Y_{t_k}\|$;
2. $\|Y_{t_{k+1}}\| = \|Y_{t_k}\| + 1$.

En el primer evento la conclusión es directa, pues por construcción $Z_{t_k} \leq Z_{t_{k+1}}$. Para el segundo evento recordemos que de la Definición 2.9 se tiene

$$\|Y_{t_{k+1}}\| = \|Y_{t_k}\| + 1 \implies M_c((t_k, t_{k+1}] \times E_c(Y_{t_k})) + M_m((t_k, t_{k+1}] \times E_m(Y_{t_k})) > 0.$$

Por otro lado, gracias a la elección de C^* en la Definición 2.13 y la hipótesis de inducción en que $\|Y_{t_k}\| \leq Z_{t_k}$, se deduce

$$E_c(Y_{t_k}), E_m(Y_{t_k}) \ \subseteq \ \{1, 2, \ldots, \|Y_{t_k}\|\} \times \mathbb{T} \times [0, C^*] \ = \ E_z(\|Y_{t_k}\|) \ \subseteq \ E_z(Z_{t_k}),$$

de donde se concluye

$$0 \ < \ M_c((t_k, t_{k+1}] \times E_z(Z_{t_k})) + M_m((t_k, t_{k+1}] \times E_z(Z_{t_k})).$$

Notemos finalmente que por definición Z no posee saltos en (t_k, t_{k+1}), y por lo tanto la positividad de la expresión anterior indica que t_{k+1} es un tiempo de salto para Z, y por lo tanto $Z_{t_{k+1}} = Z_{t_k} + 1$ de donde se concluye el resultado.

$\square$

Para procesos de nacimiento puro en $\mathbb{N}$ con tasas lineales como Z, se muestra en [7, Capítulo XVII, sección 3] que si $0 < j < k$,

$$\mathbb{P}_j(Z_t = k) \;=\; \binom{m-1}{j-1} e^{-2C^* jt}(1-e^{-2C^*t})^{m-j}, \tag{2.1}$$

igualdad que resulta fundamental para el próximo resultado, demostrado originalmente en [10].

Proposición 2.6 *[10, Teorema 3.1] Para todo $\eta \in \mathcal{A}$ y $p \geq 1$ existen $c, b > 0$ tales que para todo $t' > 0$,*

$$\mathbb{E}_\eta \left(\sup_{t\in[0,t']} \|Y_t\|^p \right) \leq ce^{bt'} < \infty.$$

Demostración. Utilizando la Proposición 2.5, se cumple $\mathbb{P}$-casi seguramente

$$\sup_{t\in[0,t']} \|Y_t\|^p \;\leq\; \sup_{t\in[0,t']} Z_t^p \;=\; Z_{t'}^p, \tag{2.2}$$

en que la igualdad se tiene pues Z_t es un proceso no decreciente. Gracias a (2.1) y (2.2) se tiene la desigualdad,

$$\mathbb{E}_\eta \left(\sup_{t\in[0,t']} \|Y_t\|^p \right) \;\leq\; \mathbb{E}_{\|\eta\|}(Z_{t'}^p) \;=\; \sum_{m=\|\eta\|}^{\infty} m^p \binom{m-1}{m-\|\eta\|} e^{-2\|\eta\|C^*t'}(1-e^{-2C^*t'})^{m-\|\eta\|},$$

y utilizando la fórmula de Stirling, para m grande podemos aproximar el coeficiente binomial en la expresión anterior como

$$\binom{m-1}{m-\|\eta\|} \;\approx\; \frac{1}{\|\eta\|!}(m-1)^{\|\eta\|-1},$$

de donde existe una constante $C > 0$ que depende de η y de p, tal que

$$\mathbb{E}_\eta \left(\sup_{t\in[0,t']} \|Y_t\|^p \right) \;\leq\; C \sum_{m=1}^{\infty} \left[(m-1)^{\|\eta\|+p-1}\right] e^{-2\|\eta\|C^*t'}(1-e^{-2C^*t'})^{m-\|\eta\|}.$$

Definiendo una variable aleatoria G con distribución geométrica de parámetro $r := e^{-2C^*t'} \leq 1$ podemos mayorar la expresión anterior por

$$\mathbb{E}(G^{\|\eta\|+\lceil p\rceil-1})Cr^{\|\eta\|-1}(1-r)^{1-\|\eta\|},$$

pero es fácil verificar que el momento de orden $k := \|\eta\| + \lceil p\rceil - 1$ de G es de la forma $P(r)/Q(r)$, en que P es un polinomio de orden k, y $Q(r) = r^k$. Como P está acotado en $[0,1]$ por una constante $A > 0$ que depende de η y p, se deduce

$$\mathbb{E}(G^{\|\eta\|+\lceil p\rceil-1})Cr^{\|\eta\|-1}(1-r)^{1-\|\eta\|} \;\leq\; ACe^{4(\|\eta\|+\lceil p\rceil-1)C^*t'},$$

luego basta tomar $c = AC$ y $b = 4(\|\eta\| + \lceil p\rceil - 1)C^*$ para concluir el resultado.

$\square$

La proposición anterior indica en particular que $\mathbb{P}$-c.s. la cantidad de individuos del proceso se mantiene acotada sobre intervalos finitos de tiempo, lo que utilizaremos para deducir que el proceso no posee explosiones.

Proposición 2.7 *Para todo $t \geq 0$ se cumple $\mathbb{P}(N_t = \infty) = 0$.*

Demostración. Comenzaremos mostrando que para $\eta \in \mathcal{A}$ se cumple $\mathbb{P}_\eta(N_t = \infty) = 0$. Para ello notemos que de la Proposición 2.6 se deduce que $\mathbb{P}_\eta(\sup_{t\in[0,t']} \|Y_t\| = \infty) = 0$, luego

$$\mathbb{P}_\eta(N_t = \infty) \; = \; \sum_{n\in\mathbb{N}} \mathbb{P}_\eta \left(N_t = \infty \; \wedge \; \sup_{t\in[0,t']} \|Y_t\| = n \right).$$

Por otro lado, sobre el evento $\{N_t = \infty\}$ se tiene

$$\sup_{t\in[0,t']} \|Y_t\| = n \iff \sup_{m\in\mathbb{N}} \|\bar{Y}_m\| = n \iff \forall m \in \mathbb{N},\; \bar{Y}_m \in \mathcal{A}_{\leq n},$$

y entonces, utilizando probabilidades totales basta con probar

$$\mathbb{P}_\eta \left(N_t = \infty \; \wedge \; \sup_{m\in\mathbb{N}} \|\bar{Y}_m\| = n \;\middle|\; \bar{Y}_1, \bar{Y}_2, \dots \right) = 0. \tag{2.3}$$

Para deducir la igualdad anterior, notemos que $N_t = \infty \iff \sum_{n=1}^{\infty} \Delta\tau_n < \infty$, y gracias a la Proposición 2.3, las variables $\Delta\tau_{m+1}\big|_{\bar{Y}_m}$ son independientes de a pares y siguen una distribución exponencial de parámetro $Q(\bar{Y}_m)$. Como Q es una función continua gracias a la Proposición 2.2, es acotada superiormente sobre el conjunto $\mathcal{A}_{\leq n}$ que es compacto, digamos, por $\overline{Q}_n$. De esta forma el evento en que la suma de las variables $\Delta\tau_n$ es finita tiene probabilidad nula, de donde se deduce (2.3) y consecuentemente $\mathbb{P}_\eta(N_t = \infty) = 0$. Como esto se cumple para todo $\eta \in \mathcal{A}$, entonces $\mathbb{P}(N_t = \infty) = 0$.

□

La proposición anterior permite deducir para cada $t \geq 0$ la existencia de un conjunto de medida nula tal que en su complemento se cumple $N_t < \infty$. Como N_t es creciente en t es sencillo verificar que existe un conjunto de medida nula tal que en su complemento se cumple $N_t < \infty$ para cualquier valor de t.

2.4.2. Absorción exponencial

Por definición la configuración vacía cumple $Q(0) = 0$ y por lo tanto una vez que el proceso llega a esta configuración se cumple que el siguiente tiempo de salto es infinito, es decir, el proceso es *absorbido*. Utilizaremos la notación $\mathcal{T}$ para referirnos al tiempo de absorción de Y. El siguiente resultado se desarrolla en detalle en [4].

Proposición 2.8 *[4, Lema 2.4] Existen θ_0 y $c : \mathbb{N} \to \mathbb{R}^+$ tales que para todo $\theta < \theta_0$ y todo $\eta \in \mathcal{A}^{-0}$*

$$\mathbb{E}_\eta(e^{\theta\mathcal{T}}) \leq c(\|\eta\|).$$

De acá se deduce directamente que $\mathbb{P}(\mathcal{T} = \infty) = 0$, esto es, fuera de un conjunto de medida nula el proceso es absorbido en tiempo finito. Podemos interpretar este resultado pensando que Y modela una población destinada a extinguirse.

2.5. Ecuación integral y propiedad de martingala

En la presente sección definiremos débilmente el generador infinitesimal del proceso Y, y mostraremos formalmente una clase de funciones para la cual este proceso cumple la propiedad de martingala. Para ello resultará fundamental mostrar que Y cumple la siguiente ecuación estocástica.

Proposición 2.9 *Sea Θ un conjunto de probabilidad nula tal que en su complemento se cumple $N_t < \infty$ para todo $t \geq 0$. Definiendo fuera de Θ el proceso I como*

$$\begin{aligned} I_t := Y_0 \; + & \int\limits_{(0,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^*} \delta_{\vec{H}_i(Y_{s-})} \mathbf{1}_{\{\, i \leq \|Y_{s-}\| \,\wedge\, \theta \,\leq\, b_{\vec{H}_i(Y_{s-})}(Y_{s-})\}} M_c(ds, di, dz, d\theta) \\ + & \int\limits_{(0,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^*} \delta_z \; \mathbf{1}_{\{\, i \leq \|Y_{s-}\| \,\wedge\, \theta \,\leq\, m_{\vec{H}_i(Y_{s-})}(Y_{s-}) g(\vec{H}_i(Y_{s-}), z)\}} M_m(ds, di, dz, d\theta) \\ - & \int\limits_{(0,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^*} \delta_{\vec{H}_i(Y_{s-})} \mathbf{1}_{\{\, i \leq \|Y_{s-}\| \,\wedge\, \theta \,\leq\, \lambda_{\vec{H}_i(Y_{s-})}(Y_{s-})\}} M_d(ds, di, d\theta), \end{aligned}$$

se cumple $Y_t = I_t$ para todo $t \geq 0$.

Demostración. Mostraremos la igualdad inductivamente.

1. Para el caso $t = \tau_0 := 0$ se tiene por definición $I_0 = Y_0(\omega)$;
2. Supongamos cierto el resultado para τ_k y sea $t \in (\tau_k, \tau_{k+1})$ cualquiera. Utilizando la hipótesis de inducción se tiene

$$\begin{aligned} I_t \;\; = \;\; Y_{\tau_k}(\omega) \; + & \int\limits_{(\tau_k,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^*} \delta_{\vec{H}_i(\bar{Y}_k)} \mathbf{1}_{\{\, i \leq \|\bar{Y}_k\| \,\wedge\, \theta \,\leq\, b_{\vec{H}_i(\bar{Y}_k)}(\bar{Y}_k)\}} M_c(ds, di, dz, d\theta) \\ + & \int\limits_{(\tau_k,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^*} \delta_z \; \mathbf{1}_{\{\, i \leq \|\bar{Y}_k\| \,\wedge\, \theta \,\leq\, m_{\vec{H}_i(\bar{Y}_k)}(\bar{Y}_k) g(\vec{H}_i(\bar{Y}_k), z)\}} M_m(ds, di, dz, d\theta) \\ - & \int\limits_{(\tau_k,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^*} \delta_{\vec{H}_i(\bar{Y}_k)} \mathbf{1}_{\{\, i \leq \|\bar{Y}_k\| \,\wedge\, \theta \,\leq\, \lambda_{\vec{H}_i(\bar{Y}_k)}(\bar{Y}_k)\}} M_d(ds, di, d\theta), \end{aligned}$$

   pues en (τ_k, τ_{k+1}) el proceso Y no presenta saltos, y por lo tanto $Y_s(\omega) = Y_{t_k}(\omega) = \bar{Y}_k(\omega)$ para s en tal intervalo. Es directo de la construcción de los conjuntos $E_x^{k,t}$ introducidos en la Definición 2.9, que la expresión anterior queda de la forma

$$\begin{aligned} I_t &= Y_{\tau_k} + \int\limits_{E_c^{k,t}} \delta_{\vec{H}_i(\bar{Y}_k)} M_c(ds, di, dz, d\theta) \\ &\quad + \int\limits_{E_m^{k,t}} \delta_z M_m(ds, di, dz, d\theta) - \int\limits_{E_d^{k,t}} \delta_{\vec{H}_i(\bar{Y}_k)} M_d(ds, di, d\theta) = Y_{\tau_k}, \end{aligned}$$

pues en (τ_k, τ_{k+1}) no ocurren saltos, luego para $x = c, m, d$ se cumple $M_x(E_x^{k,t}) = 0$. Deducimos entonces que $I_t = Y_{t_k} = Y_t$;

3. Supongamos ahora que el resultado es cierto para τ_k y veamos que se cumple también para τ_{k+1}. De forma similar al caso anterior, y recordando la definición de los conjuntos E_x^k introducidos en la Definición 2.9, se tiene

$$\begin{aligned} I_{\tau_{k+1}} &= Y_{\tau_k} + \int\limits_{E_c^k} \delta_{\vec{H}_i(\bar{Y}_k)} M_c(ds, di, dz, d\theta) \\ &\quad + \int\limits_{E_m^k} \delta_z M_m(ds, di, dz, d\theta) - \int\limits_{E_d^k} \delta_{\vec{H}_i(\bar{Y}_k)} M_d(ds, di, d\theta) \\ &= \bar{Y}_k(\omega) + v_c^k + v_m^k - v_d^k = \bar{Y}_{k+1} = Y_{t_{k+1}}(\omega), \end{aligned}$$

en que hemos utilizado la construcción de las variables $\bar{Y}_k$ y que gracias a la Proposición 2.4, son iguales al proceso Y mirado en los tiempos de salto.

□

2.5.1. Martingalas de cuadrado integrable

El propósito de expresar Y_t como una integral estocástica es deducir la propiedad de martingala del proceso para una clase bien definida de funciones. Para ello presentamos las siguientes definiciones y resultados que se tratan en detalle en [17].

Definición 2.15 *Sean $(\Omega, \mathcal{F})$ y (E^*, Σ^*) dos espacios medibles en que el primero es de probabilidad, y $\mathcal{B}_{[0,\infty)}$ la σ-álgebra de los borelianos en el intervalo $[0, \infty) \subseteq \mathbb{R}$. Para cada $t > 0$ definimos*

$$\Sigma_t := \mathcal{B}_{[0,t]} \otimes \Sigma^*,$$

y dada $M : \Omega \times (\mathcal{B}_{[0,\infty)} \otimes \Sigma^*) \to \mathbb{R}$, una medida de Poisson, se define $\mathcal{F}_t \subseteq \mathcal{F}$ como la σ-álgebra generada por las variables aleatorias

$$\{M(\cdot, B)\}_{B \in \Sigma_t}.$$

Es directo de la definición que la familia $\{\mathcal{F}_t\}_{t \geq 0}$ es una filtración continua por la derecha.

Definición 2.16 *Sean* $(\Omega, \mathcal{F})$, (E^*, Σ^*) *y* M *como antes, en que* M *tiene medida de intensidad* μ. *Se define* $q_M : \Omega \times (\mathcal{B}_{[0,\infty)} \otimes \Sigma^*) \to \mathbb{R}$, *la medida compensada de Poisson*

$$q_M := M - \mu.$$

En [17] se muestra que para un proceso simple de cuadrado integrable respecto a la medida de intensidad μ, es posible definir su integral respecto a q_M, la cual resulta ser una martingala respecto a la filtración $\{\mathcal{F}_t\}$. Para extender esta propiedad a otros procesos definimos la siguiente norma.

Definición 2.17 *Sean* $T \geq 0$ *y* $\psi : \Omega \times [0, T] \times E^* \to \mathbb{R}$ *un proceso medible. Tomando* μ *la medida de intensidad de* M *se define* $\|\psi\|_T$ *como*

$$\|\psi\|_T^2 \ := \ \mathbb{E}\left(\int_0^T \int_{E^*} \psi^2(\cdot, s, y)\mu(dy)ds\right).$$

El siguiente resultado muestra una clase suficientemente grande de procesos para los cuales su integral respecto a la medida q_M es una martingala.

Teorema 2.1 *[17, Proposición 1.21] Sea* ψ *como en la Definición 2.17, en que* $\psi(\cdot, t, y)$ *es* $\mathcal{F}_t$*-medible para cada* $y \in E^*$, $t \geq 0$. *Si* $\psi(\omega, \cdot, y)$ *es continuo por la izquierda con límite por la derecha y* $\|\psi\|_T < \infty$, *entonces el proceso* $\{int(\psi)_t\}_{t \leq T}$ *donde*

$$int(\psi)_t \ := \ \int_0^t \int_{E^*} \psi(\cdot, s, y) q_\mu(dy)ds,$$

es una martingala respecto a la filtración $\{\mathcal{F}_t\}_{t \leq T}$.

2.5.2. El generador infinitesimal

Definiremos a continuación un operador G sobre el espacio de funciones medibles en $\mathbb{R}^+ \times \mathcal{A}$ que será fundamental en el resto de este trabajo.

Definición 2.18 *Sea* $f : \mathbb{R}^+ \times \mathcal{A} \to \mathbb{R}$ *una función medible y localmente acotada con* $f(\cdot, 0) = 0$. *Se define* $G(f) : \mathbb{R}^+ \times \mathcal{A} \to \mathbb{R}$ *(que notamos como* Gf*) de la forma*

$$\begin{aligned} Gf(t, \eta) \ &= \ \sum_{y \in \{\eta\}} \eta_y b_y(\eta)[f(t, \eta + \delta_y) - f(t, \eta)] + \sum_{y \in \{\eta\}} \eta_y \lambda_y(\eta)[f(t, \eta - \delta_y) - f(t, \eta)] \\ &\quad + \sum_{y \in \{\eta\}} \eta_y m_y(\eta) \int_{\mathbb{T}} [f(t, \eta + \delta_z) - f(t, \eta)] g(y, z) d\sigma(z). \end{aligned}$$

El siguiente resultado, probado en [4], muestra que el operador G corresponde al generador débil del proceso Y. Para demostrar esta afirmación utilizaremos la notación $\mathcal{T}_n$ para referirnos al tiempo de llegada de Y al conjunto $\mathcal{A}_n$.

Proposición 2.10 *[4, Proposición 2.5] Sea $f : \mathbb{R}^+ \times \mathcal{A} \to \mathbb{R}$ medible tal que $f(\cdot, \eta)$ es continuamente diferenciable para cada η. Asumamos ciertas las hipótesis en (H) y que $f(\cdot, 0) = 0$. Definiendo el proceso $\{\mathcal{M}_t\}_{t\geq 0}$ como*

$$\mathcal{M}_t \ := \ f(t, Y_t) - f(0, Y_0) - \int_0^t \left[\frac{\partial f}{\partial x_1}(s, Y_s) + Gf(s, Y_s)\right] ds,$$

se cumplen las siguientes propiedades:

1. *Si existe una función $C : \mathbb{R}^+ \to \mathbb{R}^+$ tal que $\forall t \leq t_0$ y para $\eta \in \mathcal{A}$ cualquiera*
$$|f(t, \eta)| \leq C(t_0)(1 + \|\eta\|^{p'}),$$
en que $p' > 0$, entonces $\mathcal{M}_t$ es una martingala en $[0, t_0]$;

2. *Si f es localmente acotada, entonces para todo $n \in \mathbb{N}$, $\mathcal{M}_{\mathcal{T}_n \wedge t}$ es una martingala, luego $\mathcal{M}_t$ es una martingala local.*

Demostración. Sea $t \geq 0$ cualquiera y supongamos que Y no posee explosiones (esto ocurre $\mathbb{P}$-c.s. gracias a la Proposición 2.7). De la construcción de N_t se tiene

$$\int_0^t \frac{\partial f}{\partial x_1}(s, Y_s) ds \ = \ \sum_{k=0}^{N_t - 1} \int_{[\tau_k, \tau_{k+1})} \frac{\partial f}{\partial x_1}(s, Y_s) ds \ + \int_{[\tau_{N_t}, t)} \frac{\partial f}{\partial x_1}(s, Y_s) ds,$$

pero en cada uno de estos intervalos, Y_s es constante por definición y vale Y_{t_k}, de donde

$$\int_0^t \frac{\partial f}{\partial x_1}(s, Y_s) ds \ = \ \sum_{k=0}^{N_t - 1} [f(\tau_{k+1}, Y_{\tau_k}) - f(\tau_k, Y_{\tau_k})] \ + \ f(t, Y_t) - f(\tau_{N_t}, Y_{\tau_{N_t}}). \tag{2.4}$$

Ahora, como Y es un proceso de saltos, se tiene $Y_{\tau_{k+1}^-} := \lim_{t \nearrow \tau_{k+1}} Y_t = Y_{\tau_k}$, y entonces reemplazando (2.4) en la definición de $\mathcal{M}_t$ se tiene

$$\begin{aligned}
\mathcal{M}_t + \int_0^t Gf(s, Y_s) ds \ &= \ f(\tau_{N_t}, Y_{\tau_{N_t}}) - f(0, Y_0) - \sum_{k=0}^{N_t - 1} \left[f(\tau_{k+1}, Y_{\tau_{k+1}^-}) - f(\tau_k, Y_{\tau_k})\right] \\
&= \ \sum_{k=0}^{N_t} \left[f(\tau_{k+1}, Y_{\tau_{k+1}}) - f(\tau_{k+1}, Y_{\tau_{k+1}^-})\right].
\end{aligned}$$

Abusando entonces de la notación podemos escribir

$$\mathcal{M}_t \ + \ \int_0^t Gf(s, Y_s) ds \ = \ \sum_{s \leq t} [f(s, Y_s) - f(s, Y_{s^-})], \tag{2.5}$$

y un cálculo sencillo revela que el término de la derecha es igual a

$$\begin{aligned}
&\int_{(0,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^*} \left[f\left(s, Y_{s^-} + \delta_{\vec{H}_i(Y_{s^-})}\right) - f(s, Y_{s^-})\right] \mathbf{1}_{\{\, i \leq \|Y_{s^-}\| \,\wedge\, \theta \,\leq\, b_{\vec{H}_i(Y_{s^-})}(Y_{s^-})\}} M_c(ds, di, dz, d\theta) \\
&+ \int_{(0,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^*} [f(s, Y_{s^-} + \delta_z) - f(s, Y_{s^-})] \mathbf{1}_{\{\, i \leq \|Y_{s^-}\| \,\wedge\, \theta \,\leq\, m_{\vec{H}_i(Y_{s^-})}(Y_{s^-}) g(\vec{H}_i(Y_{s^-}), z)\}} M_m(ds, di, dz, d\theta) \\
&- \int_{(0,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^*} \left[f\left(s, Y_{s^-} - \delta_{\vec{H}_i(Y_{s^-})}\right) - f(s, Y_{s^-})\right] \mathbf{1}_{\{\, i \leq \|Y_{s^-}\| \,\wedge\, \theta \,\leq\, \lambda_{\vec{H}_i(Y_{s^-})}(Y_{s^-})\}} M_d(ds, di, d\theta).
\end{aligned}$$

Podemos definir entonces de manera natural los procesos

$$\psi_c(s,i,z,\theta) := \left[f\left(s, Y_{s^-} + \delta_{\vec{H}_i(Y_{s^-})}\right) - f(s,Y_{s^-})\right] \mathbf{1}_{\{\, i\leq \|Y_{s^-}\| \,\wedge\, \theta \leq b_{\vec{H}_i(Y_{s^-})}(Y_{s^-})\}}\,;$$

$$\psi_m(s,i,z,\theta) := [f(s, Y_{s^-} + \delta_z) - f(s,Y_{s^-})] \mathbf{1}_{\{\, i\leq \|Y_{s^-}\| \,\wedge\, \theta \leq m_{\vec{H}_i(Y_{s^-})}(Y_{s^-}) g(\vec{H}_i(Y_{s^-}),z)\}}\,;$$

$$\psi_d(s,i,z,\theta) := \left[f\left(s, Y_{s^-} - \delta_{\vec{H}_i(Y_{s^-})}\right) - f(s,Y_{s^-})\right] \mathbf{1}_{\{\, i\leq \|Y_{s^-}\| \,\wedge\, \theta \leq \lambda_{\vec{H}_i(Y_{s^-})}(Y_{s^-})\}},$$

con los cuales (2.5) queda de la forma

$$\mathcal{M}_t = \int\limits_{(0,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^*} \sum_{x=c,d,m} \psi_x(s,i,z,\theta) M_x(ds,di,dz,d\theta) \;-\; \int_0^t Gf(s,Y_s)ds. \tag{2.6}$$

Por otro lado, notemos que Y_s es igual a Y_{s^-} c.t.p. en s, luego se tiene

$$\int_0^t Gf(s,Y_s)ds \;=\; \int_0^t Gf(s,Y_{s^-})ds,$$

y un cálculo sencillo revela que

$$\begin{aligned}
Gf(s,Y_{s^-}) \;&=\; \sum_{i=1}^{\|Y_{s^-}\|} b_{\vec{H}_i(Y_{s^-})}(Y_{s^-})[f\left(s, Y_{s^-} + \delta_{\vec{H}_i(Y_{s^-})}\right) - f(s,Y_{s^-})] \\
&\quad+\; \sum_{i=1}^{\|Y_{s^-}\|} m_{\vec{H}_i(Y_{s^-})}(Y_{s^-}) \int_{\mathbb{T}} [f(s, Y_{s^-} + \delta_z) - f(s,Y_{s^-})] g(\vec{H}_i(Y_{s^-}),z) d\sigma(z) \\
&\quad+\; \sum_{i=1}^{\|Y_{s^-}\|} \lambda_{\vec{H}_i(Y_{s^-})}(Y_{s^-})[f\left(s, Y_{s^-} - \delta_{\vec{H}_i(Y_{s^-})}\right) - f(s,Y_{s^-})] \\
&=\; \int\limits_{\mathbb{N}\times\mathbb{T}\times\mathbb{R}^+} \psi_c(s,i,z,\theta) di d\sigma(z) d\theta \;+\; \int\limits_{\mathbb{N}\times\mathbb{T}\times\mathbb{R}^+} \psi_m(s,i,z,\theta) di d\sigma(z) d\theta \\
&\qquad+ \int\limits_{\mathbb{N}\times\mathbb{T}\times\mathbb{R}^+} \psi_c(s,i,z,\theta) di d\sigma(z) d\theta.
\end{aligned}$$

Reemplazando esta igualdad en (2.6) se deduce finalmente

$$\mathcal{M}_t = \int\limits_{(0,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^*} \sum_{x=c,d,m} \psi_x(s,i,z,\theta) q_x(ds,di,dz,d\theta),$$

en que q_c, q_m y q_d corresponden a las medidas compensadas de M_c, M_m y M_d respectivamente. Así, $\mathcal{M}_t$ se descompone en la integral de tres procesos respecto a sus respectivas medidas compensadas de Poisson. Tales procesos son $\mathcal{F}_t$-adaptados y continuos por la izquierda con límite por la derecha pues Y_{s^-} lo es, luego, si tales procesos tienen norma $\|\cdot\|_T$ finita podemos utilizar el Teorema 2.1 para concluir el resultado. Utilizaremos entonces las hipótesis sobre f para mostrar que los procesos ψ_x cumplen $\|\psi_x\|_T < \infty$.

1. Supongamos que existe C tal que para todo $T \geq 0$ y todo $t \leq T$

$$|f(t,\eta)|+ \leq C(T)(1+\|\eta\|^{p'}).$$

Por simplicidad mostraremos únicamente que $\|\psi_d\|_T < \infty$, pues los casos ψ_c y ψ_m son similares. Utilizando la notación

$$h(\theta,i,\eta) := \mathbf{1}_{\{i\leq\|\eta\|\}}\mathbf{1}_{\{\theta\leq\lambda_{\vec{H}_i(\eta)}(\eta)\}},$$

podemos acotar $\|\psi_d\|_T^2$ de la siguiente manera,

$$\begin{aligned}
\|\psi_d\|_T^2 &= \mathbb{E}\left(\int_{[0,T]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^+}\left[f(s,Y_{s^-}-\delta_{H^i(Y_{s^-})})-f(s,Y_{s^-})\right]^2 h(\theta,i,Y_{s^-})d\theta d\sigma(z)dids\right)\\
&\leq \mathbb{E}\left(\int_{[0,T]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^+} 2\left[f^2(s,Y_{s^-}-\delta_{H^i(Y_{s^-})})+f^2(s,Y_{s^-})\right]h(\theta,i,Y_{s^-})d\theta d\sigma(z)dids\right)\\
&\leq \mathbb{E}\left(\int_{[0,T]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^+} 16C^2(T)\|Y_{s^-}\|^{2p'}\mathbf{1}_{\{i\leq\|Y_{s^-}\|\}}\mathbf{1}_{\{\theta\leq\lambda_{\vec{H}_i(Y_{s^-})}(Y_{s^-})\}}d\theta d\sigma(z)dids\right)\\
&\leq \mathbb{E}\left(\int_{[0,T]\times\mathbb{N}} 16C^2(T)\|Y_{s^-}\|^{2p'}\mathbf{1}_{\{i\leq\|Y_{s^-}\|\}}c\|Y_{s^-}\|^p dids\right),
\end{aligned}$$

donde la última desigualdad se tiene como consecuencia de (H2). Utilizando la Proposición 2.6 podemos entonces concluir

$$\begin{aligned}
\|\psi_d\|_t^2 &\leq \mathbb{E}\left(\int_{[0,T]} 16C^2(T)c\|Y_{s^-}\|^{2p'+p+1}ds\right)\\
&\leq T16C^2(T)c\,\mathbb{E}\left(\sup_{s\in[0,T]}\|Y_s\|^{2p'+p+1}\right) < \infty,
\end{aligned}$$

de donde se sigue el resultado.

2. Mostraremos que si f es localmente acotada, entonces $\mathcal{M}_{t\wedge\mathcal{T}_n}$ es una martingala. En efecto, $\mathcal{M}_{t\wedge\mathcal{T}_n}$ corresponde a las integrales de los procesos $\mathbf{1}_{\{s\leq\mathcal{T}_n\}}\psi_x$ respecto a las medidas q_x, y así, es suficiente mostrar que la norma de estos procesos es finita. Desarrollamos únicamente el caso de ψ_d, pues los otros dos son similares.

   Recordemos que $\mathcal{A}_{\leq n}$ es un conjunto compacto, luego por hipótesis existe $M > 0$ tal que $f|_{[0,T]\times\mathcal{A}_{\leq n}} \leq M$. De acá se sigue

$$
\begin{aligned}
\|\mathbf{1}_{\{s\leq \mathcal{T}_n\}}\psi_d\|_T^2 &\leq \mathbb{E}\left(\int_{[0,T]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^+} \mathbf{1}_{\{s\leq \mathcal{T}_n\}} 4M^2 h(\theta,i,Y_{s^-})d\theta d\sigma(z)dids\right)\\
&= \mathbb{E}\left(\int_{[0,T]\times\mathbb{N}} \mathbf{1}_{\{s\leq \mathcal{T}_n\}} 4M^2 \mathbf{1}_{\{i\leq \|Y_{s^-}\|\}} \lambda_{\vec{H}_i(Y_{s^-})}(Y_{s^-})dids\right)\\
&\leq \mathbb{E}\left(\int_{[0,T]} \mathbf{1}_{\{s\leq \mathcal{T}_n\}} 4M^2 c\|Y_{s^-}\|^{p+1}ds\right)\\
&\leq T4M^2cn^{p+1} \ < \ \infty,
\end{aligned}
$$

de donde se concluye el resultado.

□

2.6. Continuidad de las tasas y extensiones de la notación

Las tasas de transición introducidas en la Definición 2.5 son consideradas continuas, sin embargo es importante notar que en ninguna parte de las demostraciones realizadas en este capítulo se utilizó la continuidad de tales tasas, sino más bien que son funciones acotadas, y por lo tanto es posible prescindir de esta hipótesis manteniendo los resultados vistos hasta el momento.

En lo que resta de este trabajo trabajaremos a menudo con procesos de saltos distintos de Y, todos ellos dentro del contexto de una población con nacimientos y muertes. Estandarizaremos entonces la notación utilizada en este capítulo de la siguiente forma.

Definición 2.19 *Sea $\{X_t\}_{t\geq 0}$ un proceso de saltos a tiempo continuo sobre un espacio de estados $\mathcal{X}$. Utilizaremos la notación*

- *τ_n^X, el n-ésimo tiempo de salto;*
- *$\mathcal{T}_B^X$, el tiempo de la primera llegada del proceso X al conjunto B;*
- *$\mathcal{T}_n^X$, el primer tiempo en que el proceso cuenta con n individuos;*
- *N_t^X, la cantidad de saltos del proceso X hasta el tiempo t.*

En general, si el contexto lo permite suprimiremos el índice X en las notaciones anteriores. Por otro lado, para una función φ localmente acotada superior e inferiormente por valores positivos sobre $\mathcal{A}$, utilizaremos la notación $\overline{\varphi}_n > 0$ y $\underline{\varphi}_n > 0$ para referirnos a las cotas para esta función sobre $\mathcal{A}_{\leq n}$.

Capítulo 3

Distribuciones cuasi-estacionarias

En el presente capítulo se realizará una breve introducción a la teoría general de distribuciones cuasi-estacionarias, también llamadas q.s.d, y estudiaremos un resultado obtenido en [4] que muestra la existencia de tales objetos para el proceso Y. Comenzaremos enunciando las principales propiedades de las q.s.d, entre las que contamos su caracterización como vectores propios de algunos operadores, y su relación con las distribuciones cuasi-límites y los límites de Yaglom. A continuación extenderemos sutilmente un resultado obtenido en [4] que permite acotar la tasa de absorción de un proceso de saltos, para finalmente concluir el capítulo con algunos resultados de existencia de distribuciones cuasi-estacionarias, entre los cuales contamos el resultado mencionado anteriormente para el proceso Y.

3.1. Teoría general

En lo que resta de esta capítulo consideraremos un proceso de Markov $\{X_t\}_{t\geq 0}$ a valores en un espacio métrico $\mathcal{X}\cup C$ dotado de la σ-álgebra $\mathcal{B}$ de los borelianos, en que C es un conjunto especial, disjunto con $\mathcal{X}$, y al cual llamamos *cementerio*. Dado el semigrupo de transición P_t del proceso, definido como

$$P_t(x, A) \;:=\; \mathbb{P}(X_t \in A \,\big|\, X_0 = x),$$

diremos que el proceso Y es un proceso absorbido en C si

$$\forall t \geq 0, \quad x \in C \implies P_t(x, \{x\}) = 1,$$

esto es, una vez que el proceso llega a un estado $x \in C$ se queda en ese estado $\mathbb{P}$-casi seguramente para siempre. Entre los procesos absorbidos, merecen especial atención aquellos procesos X que son absorbidos casi-seguramente, es decir

$$\forall x \in \mathcal{X},\ \mathbb{P}_x(\mathcal{T}_C < \infty) = 1,$$

en que hemos utilizado la notación $\mathcal{T}_C$ como en 2.19. Para los procesos absorbidos casi-seguramente, el estudio de las distribuciones estacionarias es insuficiente para deducir cualquier tipo de propiedad interesante. En efecto, utilizando la notación

$$\mathbb{P}_\pi(X_t \in A) \; := \; \int_{\mathcal{X} \cup C} \mathbb{P}_x(X_t \in A) d\pi(x)$$

para una medida de probabilidad π, deducimos que las únicas distribuciones estacionarias posibles se concentran en C, pues para $A \subseteq \mathcal{X}$

$$\pi(A) \; = \; \mathbb{P}_\pi(X_t \in A) \; \leq \; \mathbb{P}_\pi(X_t \in \mathcal{X}) \; = \; \mathbb{P}_\pi(t < \mathcal{T}_C) \; \xrightarrow[t \to \infty]{} \; 0.$$

El que π se concentre en C indica que X es absorbido, pero esto no es nada nuevo, y por lo tanto π no aporta ninguna información interesante. Algunos procesos absorbidos, sin embargo, presentan una suerte de *equilibrio* al mediano plazo, que puede entenderse como un equilibrio en el evento en que el proceso no ha sido absorbido aún. Esta noción gatilla la definición de las distribuciones cuasi-estacionarias.

Definición 3.1 *Sea $\mathcal{T}_C$ el tiempo de absorción de X en C. Diremos que una medida de probabilidad π concentrada en $\mathcal{X}$ es una distribución cuasi-estacionaria si*

$$\mathbb{P}_\pi(X_t \in A \,\big|\, \mathcal{T}_C > t) \; = \; \pi(A),$$

para todo $t \geq 0$ y todo $A \subseteq \mathcal{X}$ medible.

Así, una distribución cuasi-estacionaria representa un equilibrio para el proceso condicionado a la no-extinción. La definición de estas distribuciones para procesos a tiempo discreto es análoga.

3.1.1. Propiedades básicas

Notemos que para una q.s.d. π se cumple que para todo $t \geq 0$ y para todo $B \subseteq \mathcal{X}$

$$\mathbb{P}_\pi(X_t \in B) \; = \; \mathbb{P}_\pi(X_t \in B \; \wedge \; \mathcal{T}_C > t) \; = \; \pi(B)\mathbb{P}_\pi(\mathcal{T}_C > t),$$

y en particular, para $s \neq t$

$$\frac{\mathbb{P}_\pi(X_t \in B)}{\mathbb{P}_\pi(X_s \in B)} \; = \; \frac{\mathbb{P}_\pi(\mathcal{T}_C > t)}{\mathbb{P}_\pi(\mathcal{T}_C > s)},$$

expresión que no depende de B. Podemos deducir entonces que una q.s.d. mantiene su *forma* al evolucionar el proceso, y que va cambiando únicamente su *amplitud* de acuerdo a la distribución del tiempo de extinción. El siguiente resultado, desarrollado en [5] y [22] muestra que tal variable debe distribuirse exponencialmente

Teorema 3.1 *[5, Teorema 2.2] Sea π una q.s.d. para X, entonces existe $\theta = \theta(\pi) > 0$ tal que*

$$\mathbb{P}_\pi(\mathcal{T}_C > t) \; = \; e^{-\theta t}.$$

El parámetro θ en el teorema anterior es llamado comúnmente tasa de decaimiento o de extinción de la q.s.d. π. De este resultado se obtiene directamente que para tal q.s.d.

$$\mathbb{P}_\pi(X_t \in B) \; = \; \pi(B)e^{-\theta t},$$

y que para $s < \theta(\pi)$ se tiene $\mathbb{E}_\pi(e^{s\mathcal{T}_C}) < \infty$. En particular se deduce que $\mathbb{E}_x(e^{s\mathcal{T}_C}) < \infty$ π-casi seguramente, lo que motiva la siguiente definición.

Definición 3.2 *Para $x \in \mathcal{X}$ se define la tasa θ_x^* de decaimiento exponencial como*

$$\theta_x^* \ := \ \sup\{s \geq 0, \ \mathbb{E}_x(e^{s\mathcal{T}_C}) < \infty\}.$$

La siguiente proposición, desarrollada en [5] da cuenta de la expresión *decaimiento exponencial* escogida para θ_x^*.

Proposición 3.1 *[5, Proposición 2.4] Se tiene la igualdad*

$$\theta_x^* \ = \ \liminf_{t\to\infty} \ -\frac{1}{t}\log \mathbb{P}_x(\mathcal{T}_C > t).$$

Si además $\theta_x^* = 0$ para todo $x \in \mathcal{X}$, entonces no puede existir una q.s.d..

Esta igualdad jugará un rol fundamental en el siguiente capítulo donde, bajo ciertas condiciones de irreducibilidad, se tiene que $\theta_x^* = \theta_y^*$ casi seguramente para cada par $x, y \in \mathcal{X}$. Notemos que de la proposición anterior podemos deducir que si $\theta_x^* > 0$, entonces al comenzar el proceso en el punto x la absorción ocurre al menos con una rapidez exponencial.

En algunos procesos, al comenzar desde cualquier estado la distribución de los mismos sobre el evento en que no han sido absorbidos se acerca a un equilibrio. Este tipo de equilibrio, llamado *límite de Yaglom*, es más fuerte que el concepto de q.s.d. pues permite un estado inicial determinista para el proceso.

Definición 3.3 *Dado el proceso X, diremos que una medida de probabilidad π es,*

1. *Distribución cuasi-límite si existe una medida de probabilidad μ tal que $\forall A \subseteq \mathcal{X}$ medible*

$$\lim_{t\to\infty} \mathbb{P}_\mu(X_t \in A \mid \mathcal{T}_C > t) \ = \ \pi(A);$$

2. *límite de Yaglom si para todo $A \subseteq \mathcal{X}$ y todo $x \in \mathcal{X}$*

$$\lim_{t\to\infty} \mathbb{P}_x(X_t \in A \mid \mathcal{T}_C > t) \ = \ \pi(A).$$

Es directo notar que una q.s.d. π es siempre una distribución cuasi-límite pues basta tomar $\mu = \pi$ y entonces

$$\lim_{t\to\infty} \mathbb{P}_\pi(X_t \in A \mid \mathcal{T}_C > t) \ = \ \lim_{t\to\infty} \pi(A) \ = \ \pi(A).$$

El recíproco también es cierto, es decir, si π es una q.l.d. también es una q.s.d. y la demostración de ello puede encontrarse en [22, Proposición 1]. Tomando la medida de probabilidad $\mu = \delta_x$ para algún $x \in \mathcal{X}$, podemos deducir que un límite de Yaglom siempre es una q.l.d. y por lo tanto es una q.s.d. A diferencia del caso anterior, no toda q.s.d. es un límite de Yaglom como puede verse en [30], donde un proceso de nacimiento y muerte en $\mathbb{N}$ con tasas lineales posee una familia infinita de distribuciones cuasi-estacionarias mientras que, de existir, el límite de Yaglom es único.

El siguiente resultado, probado originalmente en [20], muestra que tomando como distribución inicial una q.s.d. no importa realmente *dónde* es absorbido el proceso en C, sino más bien *cuándo* es absorbido.

Teorema 3.2 *[5, Teorema 2.6] Sea π una q.s.d. para el proceso X, entonces las variables $\mathcal{T}_C$ y $X_{\mathcal{T}_C}$ son $\mathbb{P}_\pi$-independientes.*

En particular podemos colapsar el conjunto C a un singleton $\{0\}$ sin perder ningún tipo de información relevante para el estudio de estas distribuciones y manteniendo la propiedad de Markov del proceso X. Utilizaremos en adelante la notación $\mathcal{T}$ para referirnos a $\mathcal{T}_0$.

3.1.2. Enfoque espectral

El primer problema al comenzar a estudiar la existencia de distribuciones cuasi-estacionarias para el proceso X es que no contamos con una maquinaria adecuada para trabajar con la Definición 3.1. Es claro que tal definición encierra un concepto de *invariabilidad* respecto a un cierto operador, el cual buscamos hacer explícito a partir de la noción de vector propio. Para ello comenzamos entregando algunas definiciones y notaciones que utilizaremos en lo que resta de este trabajo.

Definición 3.4 *Dado un proceso X absorbido en 0 como antes definimos:*

1. *Los conjuntos $\mathcal{C}_b(\mathcal{X}) \subseteq \mathcal{M}_b(\mathcal{X}) \subseteq \mathcal{M}(\mathcal{X})$ de funciones $f : \mathcal{X} \to \mathbb{R}$ que son continuas y acotadas, $\mathcal{B}$-medibles y acotadas, y $\mathcal{B}$-medibles respectivamente. Utilizaremos el superíndice $+$ en estos conjuntos para referirnos al caso en que las funciones tomen valores no negativos;*
2. *para una medida μ y $f \in \mathcal{M}(\mathcal{X})$ integrable utilizaremos la notación $\langle \mu, f \rangle := \int f d\mu$;*
3. *dada $f \in \mathcal{M}(\mathcal{X})^+$ utilizaremos la notación $P_t(f) \in \mathcal{M}(\mathcal{X})$ para referirnos a la función*
$$[P_t(f)](x) := \mathbb{E}_x(f(X_t),\ \mathcal{T}_0 > t),$$
lo que permite definir el semigrupo submarkoviano $\{P_t\}_{t\geq 0}$, para el cual diremos que posee la propiedad de Feller si para todo $t \geq 0$
$$P_t(\mathcal{C}_b(\mathcal{X})) \subseteq \mathcal{C}_b(\mathcal{X});$$
4. *sea $\mathcal{M}_b(\mathcal{X}) \subseteq D$ un subespacio de $\mathcal{M}(\mathcal{X})$ y $G : D \to \mathcal{M}(\mathcal{X})$ un operador lineal. Diremos que G es un generador débil de X si para todo $f \in D$ el proceso*
$$\mathcal{M}_t := f(X_t) - f(X_0) - \int_0^t Gf(X_s)ds,$$
es una martingala. Llamaremos a D el dominio de G

Al semigrupo $\{P_t\}_{t\geq 0}$ lo llamaremos semigrupo de transición de X, para el cual, extendiendo la función f a $\mathcal{X} \cup \{0\}$ como $f(0) = 0$ en el ítem 3 obtenemos la definición usual

$$[P_t(f)](\eta) = \mathbb{E}_\eta(f(X_t),\ \mathcal{T}_0 > t) = \mathbb{E}_\eta(f(X_t),\ \mathcal{T}_0 > t) + \mathbb{E}_\eta(0,\ \mathcal{T}_0 \leq t) = \mathbb{E}_\eta(f(X_t)).$$

Utilizando las notaciones y definiciones anteriores desarrollamos el siguiente resultado, cuya demostración puede encontrarse en parte en [5].

Teorema 3.3 *Para el proceso X con semigrupo de transición $\{P_t\}_{t\geq 0}$ se tiene,*

1. *Una medida de probabilidad π es una q.s.d. ssi existe $\theta > 0$ tal que para todo $f \in \mathcal{M}_b(\mathcal{X})$ y todo $t \geq 0$*
$$\langle \pi, P_t(f) \rangle = e^{-\theta t} \langle \pi, f \rangle; \tag{3.3.i}$$
2. *[5, Lema 2.9] existe una q.s.d. π ssi existe una medida de probabilidad π' y $\theta > 0$ tal que $\forall f \in \mathcal{M}_b(\mathcal{X})$,*
$$\langle \pi', P_1(f) \rangle = e^{-\theta} \langle \pi', f \rangle; \tag{3.3.ii}$$
3. *si para una medida de probabilidad π existe $\theta > 0$ tal que $\forall f \in D$,*
$$\langle \pi, G(f) \rangle \;=\; -\theta \langle \pi, f \rangle, \tag{3.3.iii}$$
entonces π es q.s.d..

Demostración. 1. Sea π una distribución cuasi-estacionaria para el proceso X y tomemos una función simple de la forma $f = \sum_{k=1}^n a_k \mathbf{1}_{A_k}$. Se tiene
$$\mathbb{E}_\pi(f(X_t) \bigm| \mathcal{T} > t) \;=\; \sum_{k=1}^n a_k \mathbb{P}_\pi(X_t \in A_k \bigm| \mathcal{T} > t) \;=\; \sum_{k=1}^n a_k \pi(A_k) \;=\; \langle \pi, f \rangle,$$
igualdad que extendemos $\forall f \in \mathcal{M}_b(\mathcal{X})^+$ gracias al teorema de convergencia monótona. Utilizando entonces el Teorema 3.1 existe $\theta > 0$ tal que
$$\langle \pi, P_t(f) \rangle = \mathbb{E}_\pi(f(X_t) \bigm| \mathcal{T} > t) \mathbb{P}_\pi(\mathcal{T} > t) = \langle \pi, f \rangle e^{-\theta t}.$$
Supongamos por otro lado que $\forall f \in \mathcal{M}_b(\mathcal{X})^+,\ \langle \pi, P_t(f) \rangle = e^{-\theta t} \langle \pi, f \rangle$. Tomando $f = 1$ deducimos $\mathbb{P}_\pi(\mathcal{T} > t) = e^{-\theta t}$, y entonces para cada $A \in \mathcal{B}$ podemos tomar $f = \mathbf{1}_A$ con lo que concluimos
$$\mathbb{P}_\pi(X_t \in A | \mathcal{T} > t) \;=\; e^{\theta t} \mathbb{E}_\pi(f(X_t),\ \mathcal{T} > t) \;=\; \langle \pi, f \rangle \;=\; \pi(A).$$
Extendiendo esta equivalencia a las funciones en $\mathcal{M}_b(\mathcal{X})$ se deduce que π es una q.s.d..

2. Supongamos que existe una q.s.d. π, luego basta tomar $\pi' = \pi$ y $t = 1$ en el Teorema 3.1 para deducir el resultado. Para la otra implicancia sean $\theta > 0$ y π' tal que $\langle \pi', P_1(f) \rangle = e^{-\theta} \langle \pi', f \rangle$ y definamos
$$\pi := \frac{1}{\alpha} \int_0^1 e^{\theta s} \langle \pi', P_s \rangle ds,$$
en que hemos utilizado la notación $\langle \pi', P_s \rangle := \int_{\mathcal{X}} P_s(x, \cdot) d\pi'(x)$ y en que α es tal que π es medida de probabilidad. Tomando $t \in [0, 1)$ y $f \in \mathcal{M}_b$ cualesquiera se tiene
$$\begin{aligned} \langle \alpha\pi, P_t(f) \rangle \;&=\; \int_0^1 e^{\theta s} \langle \pi', P_{t+s}(f) \rangle ds \;=\; \int_0^{1-t} e^{\theta s} \langle \pi', P_{t+s}(f) \rangle ds + \int_{1-t}^1 e^{\theta s} \langle \pi', P_{t+s}(f) \rangle ds \\ &=\; \int_t^1 e^{\theta(s-t)} \langle \pi', P_s(f) \rangle ds + \int_0^t e^{\theta(s+1-t)} \langle \pi', P_{s+1}(f) \rangle ds, \end{aligned}$$

pero por hipótesis se tiene

$$e^{\theta}\langle \pi', P_{s+1}(f)\rangle \;=\; e^{\theta}\langle \pi', P_1(P_s(f))\rangle \;=\; \langle \pi', P_s(f)\rangle,$$

y entonces se concluye que $\langle \alpha\pi, P_t(f)\rangle = e^{-\theta t}\langle \pi, f\rangle$. Para el caso general $t \geq 0$ existen $n \in \mathbb{N}$ y $r \in [0,1)$ tales que $t = n + r$, luego

$$\langle \pi, P_t(f)\rangle \;=\; \langle \pi, P_{n+r}(f)\rangle \;=\; \langle \pi, P_1^{(n)}(P_r(f))\rangle \;=\; e^{-\theta n}\langle \pi, P_r(f)\rangle \;=\; e^{-\theta t}\langle \pi, f\rangle,$$

como esto se cumple para toda función $f \in \mathcal{M}_b(\mathcal{X})$ utilizamos el Teorema 3.1 para deducir que π es q.s.d..

3. Supongamos que $\langle \pi, G(f)\rangle = -\theta\langle \pi, f\rangle$ para todo $f \in D$ y mostremos que π es una distribución cuasi-estacionaria. Para ello, utilizando la definición de G y el teorema de Fubini se deduce que para todo $f \in \mathcal{M}_b(\mathcal{X}) \subseteq D$ y $t > 0$,

$$\langle \pi, P_t(f)\rangle - \langle \pi, f\rangle \;=\; \int_0^t \langle \pi, GP_s(f)\rangle ds \;=\; -\theta \int_0^t \langle \pi, P_s(f)\rangle ds.$$

Definamos la función $H : \mathbb{R}^+ \to \mathbb{R}^+$ como $H(s) := \langle \pi, P_s(f)\rangle$, que gracias a la ecuación anterior cumple

$$H(t) \;=\; H(0) + \int_0^t H(s)ds.$$

Como $\|P_s(f)\|_\infty \leq \|f\|_\infty < \infty$ para todo $s > 0$, se deduce que H es de clase $\mathcal{C}^\infty(\mathbb{R}^+)$, luego es de la forma $H(s) = H(0)e^{-\theta s}$, y entonces para todo $f \in \mathcal{C}_b(\mathcal{X})$ mostramos

$$\langle \pi, P_t(f)\rangle \;=\; e^{-\theta t}\langle \pi, f\rangle.$$

El resultado sigue de la igualdad anterior y de el Teorema 3.1.

□

Utilizando el Teorema 3.3 podemos ver que una q.s.d. π puede entenderse como un vector propio por la izquierda del semigrupo $\{P_t\}_{t\geq 0}$, de un único kernel de transición P_1, y del generador débil del proceso, G.

En el próximo capítulo mostraremos que la existencia y el comportamiento de los vectores propios por la derecha del semigrupo de transición facilitan el estudio de las propiedades de las q.s.d. del proceso, así como su unicidad en algunos casos. Para el proceso Y, sin embargo, no contamos con una expresión cerrada para el semigrupo, por lo que encontrar tales vectores directamente resulta engorroso. El siguiente resultado muestra que para encontrar estos vectores es suficiente encontrar vectores propios por la derecha para el generador débil del proceso.

Proposición 3.2 *Sea G un generador débil de X con dominio D y $f \in D$ integrable. Supongamos que existe $\theta > 0$ tal que $G(f) = -\theta f$, entonces para todo $t \geq 0$*

$$P_t(f) \;=\; e^{-\theta t} f.$$

Demostración. Como G es generador débil, de la definición de $\{P_t\}_{t\geq 0}$ y el teorema de Fubini se deduce

$$P_t(f) \;=\; f + \int_0^t P_s(G(f))ds \;=\; f - \theta \int_0^t P_s(f)ds,$$

en que para le segunda igualdad hemos usado $G(f) = -\theta f$. Similarmente a lo hecho en el teorema anterior, para $x \in \mathcal{X}$ podemos definir $H(s)$ como $H(s) = [P_s(f)](x)$, que cumple

$$H(t) \;=\; H(0) + \int_0^t H(s)ds$$

y es de clase $\mathcal{C}^\infty(\mathbb{R}^+)$, luego $[P_t(f)](x) = H(s) = H(0)e^{-\theta t} = f(x)e^{-\theta t}$.

□

3.2. Cotas para la tasa de decaimiento

Dada una q.s.d. π podemos calcular directamente su tasa de decaimiento θ como

$$\theta = -\log\left(\mathbb{P}_\pi(\mathcal{T}_0 > 1)\right)$$

Este método, sin embargo, es de poca utilidad en la práctica, donde se desconoce una expresión cerrada del kernel de transición P_1. Buscamos entonces una expresión que dependa únicamente de la q.s.d. π y parámetros conocidos del modelo, en particular aquellos que forman parte del generador débil L. Para ello asumiremos que X es un proceso de saltos en $\mathcal{X}$ de forma tal que

1. la cadena de Markov $\{\bar{X}_n\}_{n\in\mathbb{N}}$ asociada a los tiempos de salto tiene kernel de transición Q (no relacionado con el kernel utilizado en el capítulo anterior) y los tiempos de salto $\Delta\tau_n$ tienen distribución exponencial de parámetro $Q(\bar{X}_{n-1}) := Q(\bar{X}_{n-1}, \mathcal{X})$;
2. X posee la propiedad de Markov fuerte.

Utilizando la cadena $\{\bar{X}_n\}_{n\in\mathbb{N}}$ definimos los siguientes conjuntos.

Definición 3.5 *Sea $A \subseteq \mathcal{X}$ un boreliano cualquiera. Se define $\{CL_n(A)\}_{n\geq 1} \subseteq \mathcal{B}$ como*

$$CL_n(A) \;=\; \{x \in \mathcal{X},\; \mathbb{P}_x(\bar{X}_n \in A) > 0 \;\wedge\; \forall k \in \{0, 1, \ldots, n-1\},\; \mathbb{P}_x(\bar{X}_k \in A) = 0\}.$$

En el caso en que $A = \{0\}$ utilizamos la notación $CL_n(0)$

Así, cada conjunto $CL_n(A)$ corresponde a los puntos que pueden llegar en n saltos al conjunto A pero no en menos de n pasos. Notando que $\{\bar{X}_n\}$ es un proceso de Markov, sabemos que para todo conjunto A en los borelianos se cumple que el kernel $P_n(\cdot, A)$ es una función $\mathcal{B}$-medible, y por lo tanto es directo que cada $CL_n(A)$ es medible.

Los conjuntos $CL_n(0)$ buscan entregar una noción de distancia entre los elementos de $\mathcal{X}$ y el estado 0, lo que se deduce del siguiente resultado, cuya demostración se encuentra en el Anexo.

Proposición 3.3 *Sean $n, m \in \mathbb{N}$ con $n > m$. Utilizando la notación $\mathcal{T}_m := \mathcal{T}_{CL_m(0)}$ se tiene que para $x \in CL_n(0)$,*

$$\mathbb{P}_x(\mathcal{T}_{m+1} \leq \mathcal{T}_m) = 1.$$

A continuación desarrollamos una generalización sutil de [4, Lema 3.2], en que se encuentra una expresión para la tasa de decaimiento θ de una q.s.d. para el proceso Y. La demostración del resultado es análoga a la presentada en [4], salvo pequeñas diferencias.

Lema 3.1 *[4, Lema 3.2] Sea π una distribución cuasi-estacionaria en $(\mathcal{X}, \mathcal{B})$ y supongamos que para cada conjunto $CL_n(\{0\})$ existen $0 < s_n < S_n$ tales que $s_n \leq Q \leq S_n$, entonces su tasa de decaimiento θ cumple*

$$\theta \ = \int_{CL_1(0)} Q(x, \{0\}) \pi(dx).$$

Demostración. Definamos para cada $n \in \mathbb{N}$ la función $a_n(t)$ como

$$a_n(t) \ := \sup_{x \in CL_n(0)} \mathbb{P}_x(\mathcal{T} \leq t).$$

Supongamos $n \geq 2$ y tomemos $x \in CL_n(0)$ cualquiera. Gracias a la Proposición 3.3 se tiene $\mathbb{P}_x(\mathcal{T} \leq t) \ = \ \mathbb{P}_x(\mathcal{T}_2 \leq \mathcal{T} \leq t)$, y utilizando la propiedad de Markov fuerte deducimos

$$\mathbb{P}_x(\mathcal{T} \leq t) \ = \ \mathbb{E}_x(\mathbf{1}_{\{\mathcal{T}_2 \leq t\}} \mathbb{P}_{X_{\mathcal{T}_2}}(\mathcal{T} \leq t - \mathcal{T}_2)) \ \leq \ \mathbb{E}_x(\mathbb{P}_{X_{\mathcal{T}_2}}(\mathcal{T} \leq t)) \ \leq \ a_2(t) \tag{3.1}$$

Supongamos ahora que $x \in CL_2(0)$, luego por definición para llegar al estado 0 deben haber ocurrido al menos dos saltos, de donde

$$\begin{aligned} \mathbb{P}_x(\mathcal{T} \leq t) \ &\leq \ \mathbb{E}_x(\mathbf{1}_{\{\tau \leq t\}} \mathbb{P}_{\bar{X}_1}(\mathcal{T} \leq t)) \\ &= \ \mathbb{E}_x(\mathbf{1}_{\{\tau \leq t \,\wedge\, \bar{X}_1 \in CL_1(0)\}} \mathbb{P}_{\bar{X}_1}(\mathcal{T} \leq t)) \ + \ \mathbb{E}_x(\mathbf{1}_{\{\tau \leq t \,\wedge\, \bar{X}_1 \notin CL_1(0)\}} \mathbb{P}_{\bar{X}_1}(\mathcal{T} \leq t)). \end{aligned}$$

Utilizando la definición de $a_1(t)$ para acotar y el primer término, y utilizando (3.1) para acotar el segundo deducimos

$$\begin{aligned} \mathbb{P}_x(\mathcal{T} \leq t) \ &\leq \ \mathbb{E}_x(\mathbf{1}_{\{\tau \leq t\}} a_1(t)) \ + \ \mathbb{E}_x(\mathbf{1}_{\{\tau \leq t\}} a_2(t)) \\ &= \ \left(1 - e^{-Q(x)t}\right) [a_1(t) + a_2(t)], \end{aligned}$$

luego, tomando supremo sobre $CL_2(0)$ a ambos lados se tiene la desigualdad

$$a_2(t) \ \leq \ \sup_{x \in CL_2(0)} \left(1 - e^{-Q(x)t}\right) [a_1(t) + a_2(t)] \ = \ \left(1 - e^{-S_2 t}\right) [a_1(t) + a_2(t)],$$

y despejando $a_2(t)$ deducimos

$$a_2(t) \ \leq \ \left(e^{S_2 t} - 1\right) a_1(t). \tag{3.2}$$

Por otro lado, para $x \in CL_1(0)$ tenemos la desigualdad $\mathbb{P}_x(\mathcal{T}_0 \leq t) \ \leq \ \mathbb{P}_x(\tau \leq t) \ \leq \ 1 - e^{-S_1 t}$, y tomando supremo en $CL_1(0)$ se deduce

$$a_1(t) \ \leq \ 1 - e^{-S_1 t}, \tag{3.3}$$

y utilizando (3.3) para acotar (3.2) se obtiene que para $x \notin CL_1(0)$,

$$0 \;\leq\; \lim_{t\to 0} \frac{\mathbb{P}_x(\mathcal{T} \leq t)}{t} \;\leq\; \lim_{t\to 0} \frac{a_2(t)}{t} \;\leq\; \lim_{t\to 0} \frac{\left(e^{S_2 t} - 1\right)\left(1 - e^{-S_1 t}\right)}{t} \;=\; 0. \tag{3.4}$$

Para encontrar el valor de este límite cuando $x \in CL_1(0)$ notemos que

$$\mathbb{P}_x(\mathcal{T} \leq t) \;\leq\; \underbrace{\mathbb{P}_x(\bar{X}_1 = 0 \;\wedge\; \mathcal{T} \leq t)}_{p_0(t)} + \underbrace{\mathbb{P}_x(\bar{X}_1 \in CL_1(0) \;\wedge\; \mathcal{T} \leq t)}_{p_1(t)} + \underbrace{\mathbb{P}_x(\bar{X}_1 \notin CL_1(0) \;\wedge\; \mathcal{T} \leq t)}_{p_2(t)},$$

luego, podemos acotar cada término por separado como sigue;

- $p_0(t) \;=\; \mathbb{P}_x(\bar{X}_1 = 0 \;\wedge\; \mathcal{T} \leq t) \;=\; \mathbb{P}_x(\bar{X}_1 = 0 \;\wedge\; \tau \leq t) \;\leq\; (1 - e^{-S_1 t})\frac{Q(x,\{0\})}{Q(x)}$;
- utilizando nuevamente la propiedad de Markov fuerte se tiene

$$p_1(t) \;\leq\; a_1(t)\mathbb{P}_x\left(X_\tau \in CL_1(0) \;\wedge\; \tau < t\right) \;\leq\; a_1(t)\mathbb{P}_x\left(\tau < t\right) \;\leq\; \left(1 - e^{-S_1 t}\right)^2;$$

- similarmente al caso anterior, utilizando (3.2) y (3.3),

$$p_2(t) \;\leq\; a_2(t)\mathbb{P}_x\left(\tau < t\right) \;\leq\; \left(e^{S_2 t} - 1\right)\left(1 - e^{-S_1 t}\right)^2.$$

Utilizando las tres cotas anteriores se deduce

$$\mathbb{P}_x(\mathcal{T} \leq t) \;\leq\; \left(1 - e^{-S_1 t}\right)\frac{Q(x,\{0\})}{Q(x)} \;+ 2\left(e^{S_2 t} - 1\right)\left(1 - e^{-S_1 t}\right)^2, \tag{3.5}$$

y por otro lado

$$\left(1 - e^{-S_1 t}\right)\frac{Q(x,\{0\})}{Q(x)} \;=\; \mathbb{P}_x(\tau \leq t \;\wedge\; \bar{X}_1 = 0) \;\leq\; \mathbb{P}_x(\mathcal{T} \leq t), \tag{3.6}$$

luego podemos dividir (3.5) y (3.6) por $t > 0$ y tomar límite cuando $t \to 0$ de donde, para $x \in CL_1(0)$,

$$\lim_{t\to 0} \frac{\mathbb{P}_x(\mathcal{T} \leq t)}{t} \;=\; Q(x,\{0\}). \tag{3.7}$$

Para concluir el lema notemos que de (3.4) y (3.5), para todo $0 < t \leq 1$ y para todo $x \in \mathcal{X}$ se tiene

$$0 \;\leq\; \frac{\mathbb{P}_x(\mathcal{T} \leq t)}{t} \;\leq\; \frac{\left(e^{S_2 t} - 1\right)\left(1 - e^{-S_1 t}\right) + \left(1 - e^{-S_1 t}\right)}{t} \;\leq\; C$$

para alguna constante $C > 0$, luego podemos utilizar el teorema de convergencia dominada, junto con (3.4) y (3.7) para deducir

$$\lim_{t\to 0} \int_{\mathcal{X}} \frac{\mathbb{P}_x(\mathcal{T} \leq t)}{t}\pi(dx) \;=\; \int_{\mathcal{X}} \lim_{t\to 0} \frac{\mathbb{P}_x(\mathcal{T} \leq t)}{t}\pi(dx) \;=\; \int_{CL_1(0)} Q(x,\{0\})\pi(dx).$$

Falta por ver que el término de la izquierda es igual a θ, pero ello se obtiene de tomar límite cuando t tiende a cero en la igualdad

$$\int_{\mathcal{X}} \frac{\mathbb{P}_x(\mathcal{T} \leq t)}{t}\pi(dx) \;=\; \frac{1}{t}\mathbb{P}_\pi(\mathcal{T} \leq t) \;=\; \frac{1}{t}(1 - e^{-\theta t}),$$

donde hemos utilizado el Teorema 3.1 pues π es q.s.d..

$\square$

Utilizando el lema anterior, θ depende de Q y de la q.s.d. π para la cual la mayor parte de las veces no existirá una expresión cerrada. Sin embargo, es trivial encontrar una cota que dependa únicamente de Q, que asumimos conocida.

Corolario 3.1 *Bajo las condiciones del lema 3.1 se deduce que* $0 \leq \theta \leq S_1$

3.3. Resultados de existencia

Como ya se había mencionado, el estudio de procesos absorbidos se ha concentrado mayormente en difusiones y en el caso en que el espacio de estados es discreto. En estos trabajos la demostración de la existencia de distribuciones cuasi-estacionarias depende fuertemente de la estructura del proceso estudiado y por lo tanto no es posible generalizar tales razonamientos a procesos más generales. Es poco lo que se ha logrado decir en general sobre la existencia de estas distribuciones, destacando los trabajos de Tweedie [27],[28] que veremos en el siguiente capítulo, y los resultados de Collet, Martínez, Méléard, y San Martín en [4],[5] que enunciaremos en esta sección.

Los dos resultados que veremos a continuación requieren la existencia de una estructura topológica en $\mathcal{X}$ que sea preservada por el semigrupo P_t del proceso. Bajo esta mirada éstos pueden parecer muy restrictivos, sin embargo pueden ser aplicados a procesos con espacio de estados numerable, a difusiones, y en particular, al proceso Y.

3.3.1. Espacio de estados compacto

El primer resultado supone que el espacio de estados es compacto y puede ser demostrado utilizando el teorema de Krein-Rutman, que provee una maquinaria suficientemente potente como para concluir en unos cuantos pasos. Sin embargo, la demostración desarrollada a continuación es similar a lo hecho en [4], donde se utiliza un teorema de punto fijo, pues de esta manera podemos ilustrar las complicaciones que surgen al considerar un espacio de estados más general.

Proposición 3.4 *[5, Proposición 2.10] Supongamos que $\mathcal{X}$ es Hausdorff compacto. Si X posee la propiedad de Feller, entonces existe una q.s.d. π.*

Demostración. Sean $\{P_t\}_{t\geq 0}$ y $\mathcal{M}_b(\mathcal{X})$ como en la Definición 3.4, luego, dotanto a este espacio de la norma del supremo deducimos que $P_1 : \mathcal{M}_b(\mathcal{X}) \to \mathcal{M}_b(\mathcal{X})$ es una función continua y monótona, es decir

$$0 \leq f \implies 0 \leq P_1(f).$$

Gracias a la propiedad de Feller del proceso podemos restringir P_1 a $\mathcal{C}_b(\mathcal{X})$ y entonces

$$P_1 : \mathcal{C}_b(\mathcal{X}) \to \mathcal{C}_b(\mathcal{X}).$$

Sea ahora $\mathcal{C}_b^*(\mathcal{X})$ el dual de $\mathcal{C}_b(\mathcal{X})$ y definamos $P_1^* : \mathcal{C}_b^*(\mathcal{X}) \to \mathcal{C}_b^*(\mathcal{X})$ el operador adjunto de P_1, el cual tiene las siguientes propiedades;

1. como P_1 es continuo, P_1^* es continuo en la topología débil-*;
2. como P_1 es monótono, $P_1^*(v)$ es monótono para cualquier $v \in \mathcal{C}_b^*(\mathcal{X})$ monótono.

Tomando entonces el conjunto

$$\mathcal{K} \ := \ \{v \in \mathcal{C}_b^*(\mathcal{X}),\ v \text{ es monótono } \wedge\ v(\mathbf{1}_{\mathcal{X}}) = 1\},$$

que es convexo y compacto en la topología débil-*, podemos definir $S_1 : \mathcal{K} \to \mathcal{K}$ como

$$S_1(v) \ := \ \frac{P_1^*(v)}{[P_1^*(v)](\mathbf{1}_{\mathcal{X}})},$$

que es débil-* continuo. Utilizando el teorema de punto fijo de Kakutani para espacios topológicos convexos y compactos, debe existir $v \in \mathcal{C}_b^*(\mathcal{X})$ tal que $S_1(v) = v$, es decir

$$v[P_1(f)] \ = \ v[P_1(\mathbf{1}_{\mathcal{X}})] \cdot v(f) \ = \ e^{-\theta} v(f) \qquad \forall f \in \mathcal{C}_b(\mathcal{X}), \tag{3.8}$$

que es una igualdad muy similar a (3.1). Finalmente, como el espacio es Hausdorff compacto podemos utilizar el teorema de representación de Riesz para deducir que existe una medida π que representa a v, luego, utilizando el Teorema 3.3 se deduce que π es una q.s.d..

□

Notemos que en esta demostración, hemos utilizado la propiedad de Feller y las propiedades del espacio $\mathcal{X}$ únicamente para restringir el operador P_1 e identificar un elemento de $\mathcal{C}_b^*(\mathcal{X})$ con una medida. La existencia de un punto fijo se obtiene de forma general para $P_1 : \mathcal{M}_b \to \mathcal{M}_b$ aunque en este caso no es necesariamente cierto que tal punto fijo pueda representarse como una medida σ-aditiva.

3.3.2. Espacio de estados localmente compacto

Vimos en el resultado anterior que para un proceso X con la propiedad de Feller, a valores en un espacio $\mathcal{X}$ Hausdorff compacto siempre existe una distribución cuasi-estacionaria. La compacidad del espacio, sin mebargo, es una suposición muy fuerte y que no se cumple para la mayor parte de los modelos que se desean estudiar. Podemos entonces preguntarnos si es posible relajar las hipótesis de forma que $\mathcal{X}$ sea sólo localmente compacto y aún así mantener la existencia de una q.s.d.

La respuesta a esta pregunta es en general negativa, como puede verse para procesos de nacimiento y muerte en $\mathbb{N}$. Para estos procesos se demuestra en [8, Teorema 1.1] que la existencia de tales distribuciones es equivalente a que el tiempo de absorción tenga momentos exponenciales finitos, condición que no se cumple de forma general.

En la demostración anterior se encontró un punto fijo en $\mathcal{C}_b^*(\mathcal{X})$, el cual identificamos con una medida gracias al teorema de representación de Riesz. En el caso en que $\mathcal{X}$ es localmente compacto, este teorema falla, aunque sí es posible identificar este espacio como el espacio de medidas finitamente aditivas y regulares en $\mathcal{X}$ [6, Teorema IV.6.2]. De esta forma, para $\mathcal{X}$ localmente compacto siempre existe una distribución finitamente aditiva cuasi-estacionaria, y buscamos por lo tanto condiciones que permitan garantizar σ-aditividad. Enunciamos entonces el siguiente resultado, probado originalmente en [4]

Teorema 3.4 *[4, Teorema 4.2] Supongamos que $\mathcal{X}$ es un espacio de estados Hausdorff localmente compacto y que X posee la propiedad de Feller. Supongamos además que existen $c, \gamma, \beta \in \mathbb{R}^+$, $0 < \gamma < 1$ y $\varphi_0 \in \mathcal{C}(\mathcal{X})$ de forma tal que*

1. *$\varphi_0 \geq 1$ y para todo $u \geq 0$, $\varphi_0^{-1}([0, u])$ es un conjunto compacto;*
2. *$P_1(\mathbf{1}_{\mathcal{X}}) \geq c\mathbf{1}_{\mathcal{X}}$;*
3. *para todo $h \in \mathcal{C}_b(\mathcal{X})$ con $0 \leq h \leq \varphi_0$ se cumple*

$$P_1(h) \ \leq \ \gamma\varphi_0 + \beta,$$

entonces existe una q.s.d. π, la que cumple $\langle \pi, \varphi_0 \rangle < \infty$.

La demostración de este teorema sigue la misma idea de la demostración de la Proposición 3.4, restringiendo P_1^* al conjunto $\mathcal{K}_K$ de los elementos del dual que son monótonos, valen 1 en $\mathbf{1}_{\mathcal{X}}$ y que además están acotados por una cierta constante K sobre los elementos menores que φ_0.

Se muestra, utilizando las condiciones 2 y 3 que las imágenes bajo S_1 de elementos en $\mathcal{K}_K$ se mantienen en este conjunto, y como es convexo y débil-* compacto se deduce la existencia de un punto fijo para S_1 en $\mathcal{K}_K$. Finalmente, gracias a la condición 1, y utilizando el teorema de Prokhorov [1, Sección 5] se concluye que este punto fijo es en realidad una medida σ-aditiva.

3.3.3. Existencia en el caso de Y

El principal resultado obtenido en [4] consiste en la demostración del Teorema 3.4, que es de un carácter más bien general, y su aplicación al proceso Y construido en el capítulo anterior. Los siguientes resultados permiten verificar que se cumplen las hipótesis del teorema en este caso.

Proposición 3.5 *[4, Proposición 2.8] Asumiendo la continuidad de las tasas de transición b, m y λ, el proceso Y posee la propiedad de Feller.*

Lema 3.2 *[4, Lemas 4.3 y 4.6.ii] Sea $\overline{\lambda}_1$ una cota superior para λ en $\mathcal{A}_1$ como en la Definición 2.19, entonces el generador G definido en la Definición 2.18 satisface*

$$-\lambda_1 \ \leq \ G(\mathbf{1}_{\mathcal{A}^{-0}}) \ \leq \ 0,$$

y por lo tanto, para todo t se cumple

$$e^{-\lambda_1 t} \ \leq \ P_t(\mathbf{1}_{\mathcal{A}^{-0}}) \ \leq \ 1.$$

Lema 3.3 *[4, Lemas 4.5 y 4.6.iii] Sean $u, v \in \mathbb{R}$ tales que $\Gamma^* < v < u < \lambda_*$ en que Γ^* y λ_* son como en (H3). Para todo $0 < a_0 < log(u/v)$ y $0 < \gamma < 1$ existen $\beta > 0$ y $a_0 < a_1 < log(u/v)$ tales que para todo $\eta \in \mathcal{A}^{-0}$*

$$\mathbb{E}_\eta(e^{a_1\|Y_1\|},\ \mathcal{T} > 1) \ \leq \ \gamma e^{a_1\|\eta\|} + \beta.$$

El Lema 3.2 muestra que P_1 cumple la condición 2 del Teorema 3.4 tomando $c := e^{-\lambda_1}$. Por otro lado, el Lema 3.3 muestra que existe una familia infinita de funciones φ_0 que cumplen la condición 3, y que son de la forma

$$\varphi_0 \ := \ e^{a_1\|\cdot\|}.$$

Notemos finalmente que estas funciones cumplen la condición 1 del Teorema 3.4, pues $1 \leq e^{a_1\|\cdot\|}$ y los conjuntos $\left[e^{a_1\|\cdot\|}\right]^{-1}([0,u])$ son de la forma $\mathcal{A}_{\leq m}$ que son compactos. Se concluye que Y cumple las condiciones del Teorema 3.4 y por lo tanto existe una q.s.d. π con $\langle \pi, e^{a_1\|\cdot\|} \rangle < \infty$.

Capítulo 4

φ-irreducibilidad y Λ^*-clasificación

En el presente capítulo introducimos el concepto de φ-irreducibilidad para procesos de Markov absorbidos, junto con una medida específica μ que extiende σ al conjunto $\mathcal{A}$, y que resultará ser *maximal* dentro de las medidas bajo las cuales Y es irreducible. La φ-irreducibilida es de gran ayuda para probar que el semigrupo de transición de Y está bien definido sobre los espacios L^p inducidos por una clase específica de medidas, y que además tales operadores son continuos. Presentamos a continuación los principales resultados de la teoría de Λ^*-clasificación para procesos absorbidos de Markov con espacios de estados generales, la que permite clasificar tales procesos como *transientes*, *recurrentes nulos* y *recurrentes positivos* de forma análoga a lo hecho para procesos de Markov irreducibles. Las referencias que seguiremos en esta última parte corresponden principalmente a los resultados obtenidos por Breyer en [2] y Tweedie en [27], [28] y [26].

4.1. φ-irreducibilidad

Para procesos de Markov sobre espacios discretos la irreducibilidad es entendida como la capacidad del proceso de comunicar cualquier par de estados, esto es, pasar desde un estado a otro en una cantidad finita de pasos con probabilidad positiva. Para espacios generales, la probabilidad de pasar desde un estado x a un estado específico y suele ser nula, y por lo tanto es necesario ampliar esta definición para incorporar la noción de comunicación entre conjuntos. La idea detrás de la φ-irreducibilidad es que no pueden haber dos conjuntos *grandes* incomunicados, y para ello nos apoyamos en las ideas de Orey en [25] para utilizar una medida φ sobre el espacio de estados, la cual refleja precisamente la noción de tamaño para conjuntos.

Comenzamos introduciendo un operador CL, que para cada conjunto B entrega el conjunto de puntos que pueden llegar a B con probabilidad positiva, esto es, todos los estados que se *comunican* con el conjunto. Consideraremos entonces, a lo largo de este capítulo un proceso absorbido de Markov X, a valores sobre un espacio $\mathcal{X}$, dotado de una métrica bajo la cual es localmente compacto, y al que además dotamos de la σ-álgebra $\mathcal{B}$ de los borelianos. El proceso X puede ser a tiempo discreto o continuo dependiendo del contexto.

Definición 4.1 *Dado el proceso X y $B \in \mathcal{B}$ se define $CL_X(B)$ de la siguiente forma:*

- *si X es un proceso a tiempo discreto*

$$CL_X(B) \ := \ \{x \in \mathcal{X},\ \exists n \in \mathbb{N},\ \mathbb{P}_x(X_n \in B) > 0\};$$

- *si X es un proceso a tiempo continuo*

$$CL_X(B) \ := \ \left\{ x \in \mathcal{X},\ \int_0^\infty \mathbb{P}_x(X_t \in B)dt > 0 \right\}.$$

Utilizaremos la notación CL en lugar de CL_X cuando el contexto lo permita.

La principal diferencia entre el caso de tiempo discreto y el de tiempo continuo es que, mientras en el primero se pide únicamente la existencia de un tiempo posible de llegada, en el segundo se pide un conjunto de medida de Lebesgue positiva de tales tiempos.

Definición 4.2 *Dada una medida σ-finita y no nula φ sobre $\mathcal{X}$, diremos que X es φ-irreducible si $\forall B \in \mathcal{B}$*

$$\varphi(B) > 0 \implies CL(B) = \mathcal{X}.$$

Diremos además que una medida con esta propiedad es maximal si

$$\varphi(B) = 0 \implies \varphi(CL(B)) = 0.$$

Así, el proceso X será φ-irreducible si φ le asigna medida positiva a todos los conjuntos *grandes*, es decir, a los que es posible llegar desde cualquier punto. Una medida de irreducibilidad será maximal si no existen conjuntos *pequeños* con medida positiva. Es sencillo notar que toda medida de irreducibilidad es absolutamente continua respecto a una medida de irreducibilidad maximal (de ahí el nombre *maximal*), y que por lo tanto todas las medidas de irreducibilidad maximales son equivalentes.

El siguiente resultado, demostrado en [27] para procesos a tiempo discreto, y que puede extenderse para procesos a tiempo continuo, muestra que siempre existe una medida de irreducibilidad maximal.

Proposición 4.1 *[27, Lema 1.1] Supongamos que existe una medida φ no nula y σ-finita tal que el proceso X es φ-irreducible, entonces existe una medida M σ-finita y no nula tal que X es M-irreducible y M es maximal.*

Una propiedad algo más débil pero fundamental en el contexto de este trabajo para una medida es la de *concentrarse bien*, propiedad que definimos a continuación.

Definición 4.3 *Dada una medida σ-finita y no nula φ, diremos que está bien concentrada para X si $\forall B \in \mathcal{B}$ se cumple*

$$\varphi(B) = 0 \implies \varphi(CL(B)) = 0.$$

Finalmente, la siguiente propiedad permite unir la teoría desarrollada para procesos a tiempo discreto con lo hecho para procesos a tiempo continuo. En particular esto permitirá trabajar con un subconjunto numerable del semigrupo de transición:

Definición 4.4 *Dado un proceso a tiempo continuo X y una medida φ sobre $\mathcal{X}$, diremos que X es de esqueleto simultaneamente φ-irreducible, o abreviadamente φ-s.s.i. si para todo $h > 0$, el proceso $\{X_{nh}\}_{n\in\mathbb{N}}$ es φ-irreducible.*

Notemos que si X es φ-s.s.i. entonces es necesariamente φ-irreducible.

4.1.1. Propiedades del operador CL para Y

A continuación mostraremos algunas propiedades básicas del operador $CL_Y : \mathcal{B} \to \mathcal{B}$, las cuales permitirán deducir propiedades interesantes sobre el proceso y en particular sobre su semigrupo de transición.

Proposición 4.2 *Dado el proceso Y como en la definición 2.11 y dado $B \subseteq \mathcal{A}^{-0}$ medible, los siguientes conjuntos son iguales:*

1. $\left\{\eta \in \mathcal{A}^{-0},\ \exists n \in \mathbb{N},\ \mathbb{P}_\eta(\bar{Y}_n \in B) > 0\right\}$;
2. $\{\eta \in \mathcal{A}^{-0},\ \forall t > 0,\ \mathbb{P}_\eta(Y_t \in B) > 0\}$;
3. $\{\eta \in \mathcal{A}^{-0},\ \exists t > 0,\ \mathbb{P}_\eta(Y_t \in B) > 0\}$;
4. $CL(B)$.

Notemos que hemos utilizado las variables $\{\bar{Y}_n\}_{n\in\mathbb{N}}$ introducidas en la Definición 2.9.

Demostración. Mostaremos $1 \subseteq 2$ y $3 \subseteq 1$ con lo que se tendrá la igualdad, pues es trivial notar que $2 \subseteq 4 \subseteq 3$.

- $\boxed{1 \subseteq 2}$ Sea $\eta \in \mathcal{A}^{-0}$ tal que existe $n \in \mathbb{N}$ con $\mathbb{P}_\eta(\bar{Y}_n \in B) > 0$ y mostremos que para $t > 0$ cualquiera se tiene $\mathbb{P}_\eta(Y_t \in B) > 0$. Para ello recordemos que gracias a la Proposición 2.3, las variables

$$\tau_n\Big|_{\bar{Y}_0,\bar{Y}_1,\cdots,\bar{Y}_{n-1}} \quad \text{y} \quad \tau_{n+1}\Big|_{\bar{Y}_0,\bar{Y}_1,\cdots,\bar{Y}_n},$$

corresponden a la suma de n y $n+1$ variables aleatorias independientes con distribución exponencial, cada una de parámetro $Q(\bar{Y}_k)$. Por otro lado, en cada salto $\|Y\|$ puede aumentar hasta en una unidad, por lo que para cada $0 \le k \le n$, se tiene $\bar{Y}_k \in \mathcal{A}_{\le m}$ donde $m := \|\eta\| + n$. Gracias a ello, los parámetros $Q(\bar{Y}_k)$ se encuentran acotados por $\underline{Q}_m$ y $\overline{Q}_m$, introducidos en la Definición 2.19, y entonces existe $\varepsilon > 0$ tal que

$$\mathbb{P}_\eta(\tau_n \le t\ \wedge\ \tau_{n+1} > t\ |\bar{Y}_1,\cdots,\bar{Y}_n)\ \ge\ \varepsilon. \tag{4.1}$$

Condicionando sobre la cantidad de saltos de Y podemos acotar $\mathbb{P}_\eta(Y_t \in B)$ de la forma

$$\mathbb{P}_\eta(Y_t \in B) \;\geq\; \mathbb{P}_\eta(Y_t \in B,\; N_t = n) \;=\; \mathbb{P}_\eta(\bar{Y}_n \in B,\; \tau_n \leq t,\; \tau_{n+1} > t),$$

en que, condicionando esta última expresión se obtiene

$$\int\limits_{\mathcal{A}_{\leq m}^{n-1} \times B} \mathbb{P}_\eta(\tau_n \leq t,\; \tau_{n+1} - \bar{T}_n > t \,\big|\, \bar{Y}_1 \in d\eta_1, \ldots, \bar{Y}_n \in d\eta_n)\mathbb{P}_\eta(\bar{Y}_1 = d\eta_1, \ldots, \bar{Y}_n = d\eta_n).$$

Utilizando finalmente (4.1) en la expresión anterior se obtiene

$$\mathbb{P}_\eta(Y_t \in B) \;\geq\; \varepsilon \int\limits_{\mathcal{A}_{\leq m}^{n-1} \times B} \mathbb{P}_\eta(\bar{Y}_1 = d\eta_1, \ldots, \bar{Y}_n = d\eta_n) \;=\; \varepsilon\mathbb{P}_\eta(\bar{Y}_n \in B) > 0.$$

- $\boxed{3 \subseteq 1}$ Mostraremos esta inclusión probando $1^c \subseteq 3^c$. Para ello, sea $\eta \in \mathcal{A}^{-0}$ tal que para todo $n \in \mathbb{N}$, $\mathbb{P}_\eta(\bar{Y}_n \in B) = 0$ y sea $t > 0$ cualquiera. Condicionando sobre la cantidad de saltos se tiene

$$\mathbb{P}_\eta(Y_t \in B) \;=\; \sum_{k=0}^{\infty} \mathbb{P}_\eta(\bar{Y}_k \in B,\; N_t = k) \;\leq\; \sum_{k=0}^{\infty} \mathbb{P}_\eta(\bar{Y}_k \in B) \;=\; 0,$$

luego se concluye el resultado.

□

El resultado anterior formaliza la noción de que el tiempo de llegada no es un factor a considerar en la comunicación entre conjuntos. La siguiente proposición, cuya demostración se encuentra en el Anexo, permite un mejor manejo de los operadores CL y CL_n, en que estos últimos fueron introducidos en la Definición 3.5 como

$$CL_n(B) \;=\; \{\eta \in \mathcal{A}^{-0},\; \mathbb{P}_\eta(\bar{Y}_n \in B) > 0 \;\wedge\; \forall k \in \{0, 1, \ldots, n-1\},\; \mathbb{P}_\eta(\bar{Y}_k \in B) = 0\}.$$

Proposición 4.3 *Para los operadores CL y $\{CL_n\}_{n \geq 0}$ se cumple:*

1. *sean $\{B_n\}_{n\geq 0}$ medibles, entonces $CL(\cup_{n=1}^{\infty} B_n) = \cup_{n=1}^{\infty} CL(B_n)$;*
2. *sean $\{B_n\}_{n\geq 0}$ medibles, entonces $CL_1(\cup_{n=1}^{\infty} B_n) \subseteq \cup_{n=1}^{\infty} CL_1(B_n)$;*
3. *para $A \subseteq B$ medibles, $CL(A) \subseteq CL(B)$;*
4. *para $n \geq 1$ y B medible, $CL_n(B) = CL_1^{(n)}(B) - \cup_{j=1}^{n-1} CL_1^j(B)$ en que $CL_1^{(n)}$ es la composición n veces del operador CL_1;*
5. *para A medible se cumple $CL(CL(A)) = CL(A)$.*

4.1.2. La medida μ

En esta sección introduciremos la medida de probabilidad μ sobre $\mathcal{A}^{-0}$, la cual puede entenderse como una extensión de σ. Veremos que Y es μ-irreducible, que tal medida es maximal, y más aún, que μ-s.s.i. lo que permitirá utilizar los resultados de Λ^*-clasificación para este caso. Debido a que σ corresponde a una medida sobre el espacio de rasgos, su extensión μ debe medir los rasgos presentes en los elementos de $\mathcal{A}^{-0}$, por lo que será necesario realizar algunas definiciones, las cuales extienden a las introducidas en la Definición 2.2.

Definición 4.5 *A partir de los elementos introducidos en la Definición 2.2, utilizaremos las siguientes notaciones:*

- *se define el conjunto $\mathbb{N}^*$ como*
$$\mathbb{N}^* = \bigcup_{n \geq 1} \mathbb{N}^n,$$
donde consideramos que $\mathbb{N}$ no contiene al cero;

- *para cada $B \subseteq \mathcal{A}$ medible y $\vec{q} \in \mathbb{N}^*$ definimos $B_{\vec{q}} := \{\eta \in B,\ \bar{\eta} = \vec{q}\}$. Dentro de tales conjuntos cada elemento η está únicamente determinado por $\vec{\eta}$, pues a partir de $\vec{q}$ podemos saber cuántos individuos poseen cada rasgo;*

- *para cada $k \in \mathbb{N}$, $\vec{q} \in \mathbb{N}^k$ y $B \subseteq \mathcal{A}_{\vec{q}}$ medibe definimos $\overline{B} := \{\vec{\eta},\ \eta \in B\} \subseteq \mathbb{T}^k$, el conjunto de vectores de rasgos ordenados asociados a los elementos de B sin sus multiplicidades.*

Utilizando estas notaciones, para un conjunto $B \subseteq \mathcal{A}$ medible y un vector $\vec{q} \in \mathbb{N}^*$, podemos referirnos al conjunto $\overline{B_{\vec{q}}}$ como el conjunto de los vectores $\vec{\eta}$ con $\eta \in B$ y $\bar{\eta} = \vec{q}$. Esto permite definir la medida μ' de la siguiente forma.

Definición 4.6 *Se define la medida μ' sobre $\mathcal{A}^{-0}$ como*

$$\mu'(B) = \sigma^k(\overline{B}) \quad \textit{para} \quad B \subseteq \mathcal{A}_{\vec{q}} \textit{ y } \vec{q} \in \mathbb{N}^k,$$

en que σ^k corresponde a la medida producto de σ sobre $\mathbb{T}^k$. De esta forma, para cada $B \subseteq \mathcal{A}^{-0}$ medible se tendrá:

$$\mu'(B) = \sum_{k \in \mathbb{N}} \sum_{\vec{q} \in \mathbb{N}^k} \sigma^k(\overline{B_{\vec{q}}}).$$

Notemos que $\sigma^k(\mathbb{T}^k) = 1$ y para cada $n \in \mathbb{N}$ hay una cantidad finita de $\vec{q} \in \mathbb{N}^*$ tales que $\mathcal{A}_{\vec{q}} \cap \mathcal{A}_n \neq \emptyset$, luego $\mu'(\mathcal{A}_n) < \infty$. Esto permite definir finalmente la medida μ como sigue.

Definición 4.7 *Sea $\{a_k\}_{k \geq 1}$ una sucesión de reales positivos tales que $\sum_{k=0}^{\infty} a_k = 1$. Se define la medida de probabilidad μ sobre $\mathcal{A}^{-0}$ como*

$$\mu(B) = \sum_{k=1}^{\infty} \frac{a_k}{\mu'(\mathcal{A}_k)} \mu'(B \cap \mathcal{A}_k),$$

que es una medida de probabilidad equivalente a μ'.

De la definición de μ' es directo notar que para $\vec{q} \in \mathbb{N}^k$ y $B \subseteq \mathcal{A}_{\vec{q}}$ se tiene

$$\mu(B) > 0 \iff \sigma^k(\overline{B}) > 0. \tag{4.2}$$

Utilizaremos esta equivalencia para deducir el siguiente teorema, el cual indica que μ es efectivamente una medida de irreducibilidad maximal.

Teorema 4.1 *Sea $E \subseteq \mathcal{A}^{-0}$ medible, se tiene:*

1. *si $\mu(E) = 0$, entonces $\mu(CL(E)) = 0$;*
2. *si $\mu(E) > 0$, entonces $CL(E) = \mathcal{A}^{-0}$.*

DEMOSTRACIÓN. 1. Notemos que para mostrar el resultado es suficiente probar que para todo $\vec{q} \in \mathbb{N}^*$ y todo $E \subseteq \mathcal{A}_{\vec{q}}$ se tiene

$$\mu(E) = 0 \Longrightarrow \mu(CL_1(E)) = 0. \tag{4.3}$$

En efecto, supongamos cierta esta propiedad y sea $E \subseteq \mathcal{A}^{-0}$ con medida nula. Para todo $\vec{q} \in \mathbb{N}^*$ se tiene $\mu(E_{\vec{q}}) \leq \mu(E) = 0$ y entonces, utilizando (4.3) y la propiedad 2 de la Proposición 4.3 se tiene

$$\mu(CL_1(E)) \;=\; \mu\left(CL_1\left(\cup_{\vec{q}\in\mathbb{N}^*} E_{\vec{q}}\right)\right) \;\leq\; \sum_{\vec{q}\in\mathbb{N}^*} \mu(CL_1(E_{\vec{q}})) \;=\; 0.$$

Por otro lado, gracias a la propiedad 4 de la Proposición 4.3 se deduce que $\forall n \in \mathbb{N}$, $CL_n(E) = CL_{n-1}(CL_1(E))$, y entonces se puede probar inductivamente que

$$\forall n \in \mathbb{N},; \;\; \mu(E) = 0 \Longrightarrow \mu(CL_n(E)) = 0.$$

El resultado sigue de verificar que $CL(E) \;=\; \cup_{n=0}^{\infty} CL_n(E)$.

Para mostrar (4.3) supongamos que $E \subseteq \mathcal{A}_{\vec{q}}$ con $\vec{q} \in \mathbb{N}^k$, luego, de (4.2) se deduce $\sigma^k(\overline{E}) = 0$. De la estructura del proceso Y es directo verificar que $CL_1(E)$ es la unión de los siguientes conjuntos:

a) el conjunto de las configuraciones que llegan a E mediante la muerte de un individuo sin la extinción de su rasgo,

$$DS \;:=\; \{\eta \in \mathcal{A}^{-0},\; \exists j \leq \#\eta,\; \bar{\eta}_j \geq 2 \;\wedge\; \eta - \delta_{y_{\eta,j}} \in E\};$$

b) el conjunto de las configuraciones que llegan a E mediante la muerte de un individuo con la extinción de su rasgo,

$$DE \;:=\; \{\eta \in \mathcal{A}^{-0},\; \exists j \leq \#\eta,\; \bar{\eta}_j = 1 \;\wedge\; \eta - \delta_{y_{\eta,j}} \in E\};$$

c) el conjunto de las configuraciones que llegan a E mediante el nacimiento de un individuo con el mismo rasgo que el padre,

$$CB \;:=\; \{\eta \in \mathcal{A}^{-0},\; \exists j \leq \#\eta,\; \eta + \delta_{y_{\eta,j}} \in E\};$$

d) el conjunto de las configuraciones que llegan a E con probabilidad positiva mediante el nacimiento de un individuo con un rasgo nuevo,

$$MB := \left\{\eta \in \mathcal{A}^{-0},\ \sum_{y\in\{\eta\}} \eta_y m_y(\eta) \int_{\eta+\delta_z\in E} g(y,z)d\sigma(z) > 0\right\}.$$

Mostraremos a continuación que cada uno de estos conjuntos tiene medida nula según μ.

a) Sea $\eta \in DS$ cualquiera, luego existe un rasgo $z \in \{\eta\}$ presente más de una vez en η y tal que $\eta - \delta_z \in E$. Como el rasgo z está presente más de una vez en η, la muerte del individuo δ_z no produce la extinción de z, y por lo tanto la configuración $\eta' := \eta - \delta_z$ posee los mismos rasgos que η. En particular se tiene que $\vec{\eta} \in \overline{E}$.

Por otro lado, sea $\eta' \in E$. Es directo que para todo rasgo $z \in \{\eta'\}$ la configuración $\eta := \eta' + \delta_z$ presenta tal rasgo más de una vez y posee los mismos rasgos que η'. En particular, $\eta \in DS$ y por lo tanto $\vec{\eta'} = \vec{\eta} \in \overline{DS}$. Se concluye entonces que

$$\overline{E} = \overline{DS}$$

y como $\sigma^k(\overline{E}) = 0$, utilizando (4.2) se concluye $\mu(DS) = 0$.

b) Por definición, cada configuración $\eta \in DE$ posee un rasgo z presente una única vez en η y tal que $\eta - \delta_z \in E$, luego en particular, η posee $k+1$ rasgos. Es claro que el conjunto DE corresponde a la unión de los conjuntos DE_j, donde

$$DE_j := \{\eta \in \mathcal{A}^{-0},\ \bar{\eta}_j = 1 \ \wedge\ \eta - \delta_{y_{\eta,j}} \in E\},$$

es el conjunto de configuraciones tales que al ordenar sus rasgos de menor a mayor, el rasgo a desaparecer corresponde al j-ésimo. Supongamos que $j = 1$ y notemos que $\eta' \in DE_1$ ssi su primer rasgo está presente una única vez, y al desaparecer la configuración resultante pertenece a E, lo que escribimos

$$DE_1 = \{\eta + \delta_z,\ \eta \in E \ \wedge\ z < \eta_1\}.$$

De la igualdad anterior podemos deducir

$$\overline{DE_1} = \{(z, \vec{\eta}),\ \vec{\eta} \in \overline{E} \ \wedge\ z < \vec{\eta}_1\},$$

en que hemos abusado ligeramente de la notación $(z, \vec{\eta})$ para referirnos a la concatenación de z con $\vec{\eta}$. Utilizando entonces el teorema de Fubini se tiene

$$\sigma^{k+1}(\overline{DE_1}) = \int_{\overline{E}} \sigma(\{z < \vec{\eta}_1\}) d\sigma^k(\vec{\eta}) = 0,$$

pues por hipótesis $\sigma^k(\overline{E}) = 0$. En particular, gracias a (4.2) se tiene que $\mu(DE_1) = 0$, y siguiendo la misma idea para el caso $j > 1$ se puede demostrar que $\mu(DE_j) = 0$ para todo j, de donde $\mu(DE) = 0$.

c) Notemos primero que si $\vec{q}_j = 1$ para todo $1 \leq j \leq k$, entonces para las configuraciones en E los rasgos están presentes una única vez y por lo tanto $CB = \emptyset$, de donde el resultado es directo. Supongamos entonces que existe j_0 tal que $\vec{q}_{j_0} > 1$ y tomemos $\eta \in CB$, luego por definición existe un rasgo z presente en η tal que $\eta + \delta_z \in E$. Como no se ha agregado ningún rasgo nuevo se tiene $\{\eta\} = \{\eta + \delta_z\}$ y por lo tanto $\overline{CB} \subseteq \overline{E}$.

   Consideremos ahora $\eta \in E \subseteq \mathcal{A}_{\vec{q}}$ cualquiera, para el cual se debe tener $\bar{\eta}_j = \vec{q}_j > 1$ y entonces, tomando z su rasgo j-ésimo se cumple $\eta - \delta_z \in CB$ y $\{\eta\} = \{\eta - \delta_z\}$ pues la muerte de δ_z no elimina tal rasgo. Se tiene entonces

$$\overline{E} \;=\; \overline{CB},$$

   y el resultado se concluye como en a).

d) Comencemos definiendo para $\eta \in MB$ cualquiera el conjunto $F_1(\eta)$ como

$$F_1(\eta) \;:=\; \{z \in \mathbb{T},\ \eta + \delta_z \in E \ \wedge\ z < \eta_1\},$$

   el conjunto de rasgos menores que los presentes en η tales que al añadirlos a esta configuración se obtiene un elemento de E. Similarmente definimos los conjuntos $F_j(\eta)$ para el caso $\vec{\eta}_{j-1} < z < \vec{\eta}_j$, y $F_k(\eta)$ para $\vec{\eta}_{k-1} < z$. Utilizando entonces estos conjuntos definimos

$$MB_j \;:=\; \left\{ \eta \in \mathcal{A}^{-0},\ \sum_{y \in \{\eta\}} \eta_y m_y(\eta) \int\limits_{F_j(\eta)} g(y,z) d\sigma(z) > 0 \right\},$$

   cuya unión es claramente MB. Estos conjuntos representan a los η que llegan a E con probabilidad positiva introduciendo un rasgo nuevo, el cual debe cumplir la relación de orden indicada por j, y por lo tanto, si $\vec{q}_j > 1$ se debe tener $MB_j = \emptyset$. Por simplicidad trabajaremos con $j = 1$ asumiendo $\vec{q}_1 = 1$, luego por definición

$$\eta \in MB_1 \ \wedge\ z \in F_1(\eta) \implies \eta + \delta_z \in E,$$

   y en particular abusando de la notación como en b), se tiene

$$\{(z, \vec{\eta}),\ z \in F_1(\eta) \ \wedge\ \vec{\eta} \in \overline{MB_1}\} \;\subseteq\; \overline{E}.$$

   Evaluando σ^k sobre estos conjuntos y utilizando el teorema de Fubini deducimos

$$\int_{\overline{MB_1}} \sigma[F_1(\eta)] d\sigma^{k-1}(\vec{\eta}) \;\leq\; \sigma^k(\overline{E}) \;=\; 0, \tag{4.4}$$

   pero por construcción, sobre MB_1 se tiene

$$\sum_{y \in \{\eta\}} \eta_y m_y(\eta) \int\limits_{F_1(\eta)} g(y,z) d\sigma(z) > 0,$$

   y en particular $\sigma(F_1(\eta)) > 0$, por lo que (4.4) indica $\sigma^{k-1}(\overline{MB_1}) = 0$. El caso $j > 1$ es análogo y por lo tanto, utilizando (4.2) se tiene $\mu(MB) = 0$.

Hemos probado entonces que DS, DE, CB y MB son de medida nula, y por lo tanto $CL_1(E)$ es de medida nula, de donde se sigue (4.3).

2. Para probar $\mu(E) > 0 \Longrightarrow CL(E) = \mathcal{A}^{-0}$ es suficiente con mostrar las siguientes propiedades:

   i) para todo $A \subseteq \mathcal{A}_1$ con $\mu(A) > 0$ se tiene $\mathcal{A}_1 \subseteq CL(A)$;

   ii) $CL(\mathcal{A}_1) = \mathcal{A}^{-0}$;

   iii) para todo $A \subseteq \mathcal{A}^{-0}$ con $\mu(A) > 0$ existe $B \subseteq \mathcal{A}_1$ tal que $B \subseteq CL(A)$ y $\mu(B) > 0$.

   En efecto, como $\mu(E) = 0$ existe $n \in \mathbb{N}$ tal que $E_n := E \cap \mathcal{A}_n$ tiene medida positiva. Gracias a $iii)$ existe $E_1' \subseteq \mathcal{A}_1$ medible tal que

   $$\mu(E_1') > 0 \;\text{ y }\; E_1' \subseteq CL(E),$$

   luego aplicando $i)$ al conjunto $E_1' \subseteq \mathcal{A}_1$, que cumple $\mu(E_1') > 0$ se tiene

   $$\mathcal{A}_1 \subseteq CL(E_1') \subseteq CL(CL(E)) = CL(E),$$

   y gracias a $ii)$ deducimos

   $$\mathcal{A}^{-0} = CL(\mathcal{A}_1) \subseteq CL(CL(E)) = CL(E),$$

   de donde se tiene el resultado deseado.

   Para concluir la demostración probaremos a continuación $i)$, $ii)$ y $iii)$:

   i) Sea $y \in \mathbb{T}$ cualquiera, y tomemos $\eta = \delta_y$. Utilizando la Proposición 4.2 es suficiente con probar $\mathbb{P}_\eta(\bar{Y}_2 \in A) > 0$, pero de las Definiciones 2.9 y 4.5 es directo que

   $$\begin{aligned} \mathbb{P}_\eta(\bar{Y}_2 \in A) &\geq \mathbb{P}_\eta(\bar{Y}_2 \in A \;\wedge\; \bar{Y}_1 \in \eta + A) \\ &= \int_A \mathbb{P}_{\eta+\delta_z}(\bar{Y}_1 \in A)\mathbb{P}_\eta(\bar{Y}_1 \in \eta + d\eta') \\ &= \int_{\overline{A}} \frac{\lambda_y(\eta+\delta_z)}{Q(\eta+\delta_z)} \frac{m_y(\eta)}{Q(\eta+\delta_z)} g(y,z) d\sigma(z), \end{aligned}$$

   luego, utilizando las cotas introducidas en la Definición 2.19 para las tasas de transición y tomando $0 < \underline{g} \leq g$ se tiene

   $$\mathbb{P}_\eta(\bar{Y}_2 \in A) \;\geq\; \sigma(\overline{A})\underline{\lambda}_2\underline{m}_2\underline{g}(\overline{Q}_2)^{-2} > 0,$$

   pues como $\mu(A) > 0$, de (4.5) se deduce $\sigma(\overline{A}) > 0$.

ii) Este resultado es directo de la propiedad de Markov y de que para todo $\eta \in \mathcal{A}^{-0}$, $\mathbb{P}_\eta(\|\bar{Y}_1\| = \|\eta\| - 1) > 0$.

iii) Sea $A \subseteq \mathcal{A}_n$ con $\mu(A) > 0$ y $n \geq 2$. Como los conjuntos $\mathcal{A}_{\vec{q}}$ particionan $\mathcal{A}$, existen $k \in \mathbb{N}$ y $\vec{q} \in \mathbb{N}^k$ tales que $\mu(A_{\vec{q}}) > 0$. Si $\vec{q}_1 > 1$, entonces el conjunto CB estudiado en c) es no vacío y cumple

$$\sigma^k(\overline{CB}) \;=\; \sigma^k(\overline{A}) \;>\; 0.$$

Se tiene entonces que existe un conjunto $CB \subseteq \mathcal{A}_{n-1}$ tal que

$$\mu(CB) > 0 \;\text{ y }\; CB \subseteq CL_1(A) \subseteq CL(A).$$

Supongamos ahora que $\vec{q}_1 = 1$, y consideremos los conjuntos MB_1 y $F_1(\eta)$ introducidos en d). Si definimos el conjunto

$$MP \;:=\; \{\eta - \delta_{\vec{\eta}_1},\; \eta \in A\},$$

es sencillo verificar que para todo $\eta \in MP$ y $z \in F_1(\eta)$ se tiene $\eta + \delta_z \in A$, y en particular

$$\{(z, \vec{\eta}),\; \vec{\eta} \in \overline{MP} \;\wedge\; z \in F_1(\eta)\} \;=\; \overline{A}.$$

Utilizando entonces el teorema de Fubini y la definición de MB_1 se tiene

$$\int_{\overline{MB_1}} \sigma(F_1(\eta)) d\sigma^{k-1}(\vec{\eta}) \;=\; \int_{\overline{MP}} \sigma(F_1(\eta)) d\sigma^{k-1}(\vec{\eta}) \;=\; \sigma^k(\overline{A}) > 0,$$

de donde en particular, $\sigma^{k-1}(MB_1)$. Nuevamente hemos encontrado un conjunto $MB_1 \subseteq \mathcal{A}_{n-1}$ tal que

$$\mu(MB_1) > 0 \;\text{ y }\; MB_1 \subseteq CL_1(A) \subseteq CL(A).$$

Hemos probado entonces que independientemente de $\vec{q}$, para $A \subseteq \mathcal{A}_n$ existe $B \subseteq \mathcal{A}_{n-1}$ de medida positiva tal que $B \subseteq CL(A)$. Utilizando finalmente las propiedades 3 y 5 de la Proposición 4.3 junto con el argumento anterior $n-1$ veces, para $A \subseteq \mathcal{A}_n$ existe $B \subseteq \mathcal{A}_1$ de medida positiva tal que $B \subseteq CL(A)$.

□

Es interesante notar que la demostración anterior muestra la complejidad de la estructura del proceso Y así como la diversidad de los eventos que pueden ocurrir. Un estudio de las propiedades de las q.s.d. basado en la estructura de Y resulta por lo tanto de una enorme dificultad y es entonces necesario recurrir a otros métodos más accesibles.

El siguiente resultado es consecuencia directa de la Proposición 4.2 e indica que la irreducibilidad de Y es equivalente a la del proceso $\{Y_n\}_{n\in\mathbb{N}}$.

Corolario 4.1 *El proceso Y es μ-s.s.i.*

Demostración. Sea $h > 0$ cualquiera y tomemos la cadena $\{Y_{nh}\}_{n\in\mathbb{N}}$, para la cual el operador CL está definido en la Definición 4.1 como

$$CL_{Y_h}(B) \ := \ \{\eta \in \mathcal{A}^{-0}, \ \exists n \in \mathbb{N}, \ \mathbb{P}_\eta(Y_{nh} \in B) > 0\}.$$

Utilizando la Proposición 4.2 se tiene que para $B \subseteq \mathcal{A}^{-0}$ medible,

$$\begin{aligned}
\eta \in CL_{Y_h}(B) \ &\Longrightarrow \ \exists n \in \mathbb{N}, \ \mathbb{P}_\eta(Y_{nh} \in B) > 0 \\
&\Longrightarrow \ \eta \in CL_Y(B) \\
&\Longrightarrow \ \forall t > 0, \ \mathbb{P}_\eta(Y_t \in B) > 0 \\
&\Longrightarrow \ \exists n \in \mathbb{N}, \ \mathbb{P}_\eta(Y_{nh} \in B) > 0 \\
&\Longrightarrow \ \eta \in CL_{Y_h}(B)
\end{aligned}$$

Y utilizando el teorema anterior tenemos $\mu(B) \implies CL_{Y_h}(B) = CL_Y(B) = \mathcal{A}^{-0}$ de donde se concluye el resultado.

□

En el capítulo anterior vimos que una q.s.d. puede entenderse como un vector propio por la izquierda del semigrupo de transición del proceso, lo que queda explícito definiendo el semigrupo de operadores adjuntos.

Definición 4.8 *Dados $t \geq 0$ y una medida ν sobre $\mathcal{A}^{-0}$ se define el dual de P_t actuando sobre ν como*

$$P_t^*(\nu)(\cdot) \ := \ \int_{\mathcal{A}^{-0}} \mathbb{P}_\eta(Y_t \in \cdot, \ \mathcal{T} > t) d\nu(\eta).$$

De esta definición y el Teorema 3.3 es directo ver que para una q.s.d. π existe $\theta > 0$ con

$$P_t^*(\pi) = e^{-\theta t}\pi. \tag{4.5}$$

Utilizaremos estos operadores adjuntos, junto con (4.5) para probar la siguiente proposición, la cual muestra una relación entre las q.s.d. y μ.

Proposición 4.4 *Sea π una q.s.d. y μ como en la Definición 4.7, entonces $\mu << \pi$. En particular, si $\pi = f d\mu$ para algún $f \geq 0$ entonces $f > 0$ μ-c.s..*

Demostración. Sea ν una medida positiva distinta de cero sobre $\mathcal{A}^{-0}$ y mostremos que

$$\mu << P_t^*(\nu). \tag{4.6}$$

Para ello, sea $E \subseteq \mathcal{A}^{-0}$ medible con $\mu(E) > 0$ de donde, gracias al Teorema 4.1 se tiene $CL(E) = \mathcal{A}^{-0}$ y entonces, para todo $\eta \in \mathcal{A}^{-0}$ y $t > 0$ se tiene

$$\mathbb{P}_\eta(Y_t \in E, \ \mathcal{T} > 0) = \mathbb{P}_\eta(Y_t \in E) > 0,$$

de donde se deduce (4.6) pues

$$P_t^*(\nu)(E) \ = \ \int_{\mathcal{A}^{-0}} \mathbb{P}_\eta(Y_t \in E, \ \mathcal{T} > 0) d\nu(\eta) > 0.$$

Si consideramos ahora una q.s.d. π en (4.6), gracias a (4.5) se tiene

$$\mu << e^{-\theta t}\pi$$

para algún $\theta > 0$, y por lo tanto se tiene $\mu << \pi$. Como consecuencia, si $\pi << \mu$, ambas medidas son equivalentes y la densidad de Radon-Nikodym debe ser positiva casi-seguramente.

□

4.2. Medidas bien concentradas

Las medidas bien concentradas, introducidas en la Definición 4.3, extienden a las medidas de irreducibilidad maximales en el sentido de que los puntos que pueden llegar a un conjunto *pequeño* forman otro conjunto *pequeño*, y su importancia puede apreciarse en el siguiente resultado.

Proposición 4.5 *Sea $t > 0$ y consideremos el operador P_t y su dual P_t^*. Dada una medida bien concentrada ν se cumple;*

1. *Si f es una función nula ν-c.s. entonces $P_t(f) = 0$, ν-c.s.*
2. *Si ν' es una medida con $\nu' << \nu$, entonces $P_t^*(\nu') << \nu$.*

Demostración. 1. Sean $N := \{\eta \in \mathcal{A}^{-0},\ f(\eta) \neq 0\}$, que es de medida nula por hipótesis, luego por definición de medida bien concentrada se deduce $\nu(CL(N)) = 0$. Probaremos el resultado mostrando que fuera de $CL(N)$ la función $P_t(f)$ es nula, y para ello notamos que dado $\eta \notin CL(N)$ se cumple

$$\begin{aligned} P_t(f)(\eta) &= \mathbb{E}_\eta(f(Y_t),\ \mathcal{T} > t) \\ &= \int_{N^c} f(\eta')\mathbb{P}_\eta(Y_t = d\eta') + \int_N f(\eta')\mathbb{P}_\eta(Y_t = d\eta'), \end{aligned}$$

en que N^c corresponde al complemento de N. Por definición de N^c, $f(\eta') = 0$ sobre este conjunto y entonces el primer término en la expresión anterior es nulo. Similarmente, como $\eta \notin CL(N)$ y gracias a la Proposición 4.2 se cumple $\mathbb{P}_\eta(Y_t \in N) = 0$ y entonces el segundo término también es nulo.

2. Sea $E \subseteq \mathcal{A}^{-0}$ medible con $\nu(E) = 0$, luego, como ν está bien concentrada se tiene $\nu(CL(E)) = 0$. Por otro lado, de la definición de P_t^* se tiene

$$P_t^*(\nu')(E) = \int_{CL(E)} \mathbb{P}_\eta(Y_t \in E)d\nu'(\eta) + \int_{CL(E)^c} \mathbb{P}_\eta(Y_t \in E)d\nu'(\eta),$$

y como $\nu(CL(E)) = 0$ y $\nu << \mu$ se tiene $\nu'(CL(E)) = 0$ de donde el primer término de la expresión anterior es nulo. Por otro lado, sobre $CL(E)^c$ se tiene $\mathbb{P}_\eta(Y_t \in E) = 0$ y entonces el segundo término también es nulo, de donde se concluye $P_t^*(\nu) << \mu$.

□

Notemos que gracias a la primera propiedad en la proposición podemos deducir la importancia de las medidas bien concentradas como una clase de medidas para la cual los operadores P_t están bien definidos c.s..

Proposición 4.6 *Sea ν una medida bien concentrada en $\mathcal{A}^{-0}$, entonces para todo t, P_t está bien definido sobre $L^p(\nu)$.*

De su definición es directo notar que una medida de irreducibilidad maximal está particularmente bien concentrada, y gracias al resultado que mostraremos a continuación, cualquier q.s.d. π posee esta propiedad.

Proposición 4.7 *Sea π una q.s.d. sobre $\mathcal{A}^{-0}$, entonces π está bien concentrada.*

Demostración. Comencemos notando notemos que para $A \subseteq \mathcal{A}^{-0}$ medible se tiene

$$signo(P_t(\mathbf{1}_A)) \ = \ \mathbf{1}_{CL(A)},$$

en que la función $signo : \mathbb{R} \to \{-1, 0, 1\}$ es nula en cero e igual al signo de su argumento en otro caso, y en que para B medible, $\mathbf{1}_B$ corresponde a la función indicatriz de B. En efecto, $signo(P_t(\mathbf{1}_A)) \geq 0$ pues $P_t(\mathbf{1}_A) \geq 0$ y entonces

$$\begin{aligned} signo(P_t(\mathbf{1}_A))(\eta) = 1 \iff & P_t(\mathbf{1}_A)(\eta) > 0 \\ \iff & \mathbb{P}_\eta(Y_t \in A) > 0 \\ \iff & \eta \in CL(A), \end{aligned}$$

en que esta última equivalencia se tiene gracias a la Proposición 4.2. De esta forma, tomando $A \subseteq \mathcal{A}^{-0}$ medible con $\pi(CL(A)) > 0$ se deduce

$$0 < \pi(CL(A)) = \langle \pi, signo(P_t(\mathbf{1}_A)) \rangle \implies 0 < \langle \pi, P_t(\mathbf{1}_A) \rangle,$$

y utilizando el Teorema 3.3 sabemos que como π es una q.s.d. existe $\theta > 0$ con

$$0 < e^{\theta t} \langle \pi, P_t(\mathbf{1}_A) \rangle = \pi(A),$$

luego se concluye $\pi(CL(A)) > 0 \Rightarrow \pi(A) > 0$ y entonces π está bien concentrada.

□

4.2.1. Continuidad del semigrupo

Gracias a la Proposición 4.6 hemos visto que los operadores P_t están bien definidos sobre los espacios $L^p(\nu)$ para medidas ν bien concentradas. Nos preguntamos entonces si para $f \in L^p(\nu)$ la función $P_t(f)$ pertenece al mismo espacio, y si es una aplicación continua. Este resultado no es cierto en general, sin embargo, añadiendo una hipótesis de monotonía para ν no sólo se tienen estas propiedades sino que además P_t es débil-* continuo como operador en $L^\infty(\nu)$.

Teorema 4.2 *Sea X un proceso absorbido a valores sobre $\mathcal{X}$ y ν una medida de probabilidad bien concentrada sobre $\mathcal{X}$, entonces para todo $t > 0$ el operador $P_t : L^\infty(\nu) \to L^\infty(\nu)$ está bien definido, es continuo y es débil-* continuo. Si además*

- $\inf_{y\in\mathcal{X}} P_t(y,\mathcal{X}) > 0$;
- *el operador CL cumple la Proposición 4.2 (para $\mathcal{X}$ en lugar de $\mathcal{A}^{-0}$);*
- *existe $\theta > 0$ tal que*
$$P_t^*(\nu) \;\leq\; e^{-\theta t}\nu,$$

entonces para cualquier $1 \leq p \leq \infty$,

$$P_t \;: L^p(\nu) \longrightarrow L^p(\nu)$$

está bien definido y es un operador continuo.

Demostración. Sea $f \in L^p(\nu)$ y supongamos primero que $p = \infty$. En este caso, sabemos que fuera de un conjunto de medida nula N existe una constante c positiva tal que $f \leq c$. Como ν está bien concentrada, $\nu(CL(E)) = 0$, y podemos utilizar un criterio similar al utilizado en la Proposición 4.5 para deducir

$$P_t(f)(y) \;=\; \int_{\mathcal{X}} f(z)P_t(y,dz) \;=\; \int_E f(z)P_t(y,dz) + \int_{E^c} f(z)P_t(y,dz) \;\leq\; c,$$

es decir, P_t es un operador lineal continuo en $L^\infty(\nu)$ con constante 1.

Para ver que P_t es débil-* continuo recordemos que una sucesión $\{f_n\}_{n\in\mathbb{N}} \subseteq L^\infty(\nu)$ converge en la topología débil-* a f (que notaremos $f_n \overset{*}{\rightharpoonup} f$) si y sólo si para toda función $g \in L^1(\nu)$ se cumple

$$\int_{\mathcal{X}} g(x)f_n(x)d\nu(x) \;\xrightarrow[n\to\infty]{}\; \int_{\mathcal{X}} g(x)f(x)d\nu(x).$$

Supongamos entonces que $f_n \overset{*}{\rightharpoonup} f$ y mostremos que $P_t(f_n) \overset{*}{\rightharpoonup} P_t(f)$. Para ello, sea $g \geq 0$ con $g \in L^1(\nu)$ y definamos la medida M como

$$M(A) \;:=\; \int_A g(x)d\nu(x),$$

que es absolutamente continua respecto a ν. Como ν está bien concentrada podemos utilizar la segunda propiedad de la Proposición 4.5 para deducir que

$$M' \;:=\; P_t^*(M) \;<<\; \nu,$$

luego existe una función h medible tal que $dM' = hd\nu$. Es directo notar que h debe ser no negativa, y como g es integrable respecto a ν, M es finita y por lo tanto h es integrable respecto a ν. De esta forma tenemos que

$$\int_{\mathcal{X}} g(x)P_t(f_n)(x)d\nu(x) \;=\; \langle P_t(f_n), M\rangle \;=\; \langle f_n, M'\rangle \;=\; \int_{\mathcal{X}} f_n(x)h(x)d\nu(x),$$

y como h pertenece a $L^1(\nu)$, por hipótesis se cumple

$$\int_{\mathcal{X}} g(x)P_t(f_n)(x)d\nu(x) \;=\; \int_{\mathcal{X}} h(x)f_n(x)d\nu(x) \;\xrightarrow[n\to\infty]{}\; \int_{\mathcal{X}} h(x)f(x)d\nu(x) \;=\; \int_{\mathcal{X}} g(x)P_t(f)(x)d\nu(x).$$

Dado ahora $g \in L^1(\nu)$ cualquiera, se deduce el mismo resultado separando g en su parte positiva y negativa, de donde se concluye que P_t es débil-* continua.

Sea ahora $f \in L^p(\nu)$ cualquiera, en que $1 \leq p < \infty$. y $t \geq 0$. Para mostrar que la función $P_t(f)$ pertenece a $L^p(\nu)$ recordemos que

$$P_t(f) \in L^p(\nu) \iff \int_{\mathcal{X}} [P_t(f)(y)]^p \, d\nu(y) < \infty,$$

y utilizando la desigualdad de Jensen se tiene

$$\begin{aligned}
\int_{\mathcal{X}} [P_t(f)(y)]^p \, d\nu(y) &= \int_{\mathcal{X}} \left(\int_{\mathcal{X}} f(z)P_t(y,dz) \right)^p d\nu(y) \\
&\leq \int_{\mathcal{X}} \int_{\mathcal{X}} f^p(z) \frac{P_t(y,dz)}{P_t(y,\mathcal{X})} d\nu(y) \\
&\leq \left(\inf_{y\in\mathcal{X}} P_t(y,\mathcal{X}) \right)^{-1} \langle P_t(f^p), \nu \rangle \\
&\leq \left(\inf_{y\in\mathcal{X}} P_t(y,\mathcal{X}) \right)^{-1} e^{-\theta t} \langle f^p, \nu \rangle,
\end{aligned}$$

donde el primer factor es finito por hipótesis y el último también pues $f \in L^p(\nu)$. De esta forma no sólo hemos mostrado que $P_t(f) \in L^p(\nu)$ sino que además P_t es una aplicación continua sobre este espacio con constante $(\inf_{y\in\mathcal{X}} P_t(y,\mathcal{X}))^{-1} e^{-\theta t}$.

□

Notemos que para el proceso Y se cumple la primera hipótesis del teorema pues

$$e^{-\overline{Q}_1 t} \;\leq\; \inf_{\eta\in\mathcal{A}_1} e^{-Q(\eta)t} \;=\; \inf_{\eta\in\mathcal{A}_1} \mathbb{P}_\eta(\tau_1 > t) \;\leq\; \inf_{\eta\in\mathcal{A}_1} P_t(\eta, \mathcal{A}^{-0}) \;\leq\; \inf_{\eta\in\mathcal{A}^{-0}} P_t(\eta, \mathcal{A}^{-0}),$$

y el primer término es positivo. De esta forma, en el contexto del proceso Y el operador P_t resultará continuo entre los espacios L^p si la medida está bien concentrada y cumple

$$P_t^*(\nu) \;\leq\; e^{-\theta t}\nu,$$

para algún $\theta > 0$, propiedad que trabajaremos más detalladamente en las siguientes secciones. Es directo de la Proposición 4.7 que las distribuciones cuasi-estacionarias cumplen ambas hipótesis, sin embargo nos gustaría encontrar una medida en concreto que también cumpla con estas hipótesis. Construiremos entonces μ_s a partir de μ de la siguiente forma.

Definición 4.9 *Dada la medida μ de la Definición 4.7 se define μ_s como*

$$\mu_s(\cdot) \;:=\; \int_{\mathcal{A}^{-0}} \int_0^\infty \mathbb{P}_\eta(Y_t \in \cdot) dt d\mu(\eta).$$

Proposición 4.8 *Existe una sucesión* $\{a_k\}_{k\in\mathbb{N}}$ *tal que para* μ *construido a partir de ella como en la Definición 4.7, la medida* μ_s *es de probabilidad, equivalente a* μ*, y para todo* $t \geq 0$*,*

$$P_t^*(\mu_s) \ \leq \ \mu_s. \tag{4.7}$$

Demostración. Gracias a la Proposición 2.8 existen $\theta > 0$ y $c : \mathbb{N} \to \mathbb{R}^+$ tales que $\forall \eta \in \mathcal{A}^{-0}$

$$\int_0^\infty \mathbb{P}_\eta(Y_t \in \mathcal{A}^{-0})dt \ = \ \mathbb{E}_\eta(\mathcal{T}) \ \leq \ \theta^{-1}\left[\mathbb{E}_\eta\left(e^{\theta\mathcal{T}}\right) - 1\right] \ \leq \ \theta^{-1}\left[c(\|\eta\|) - 1\right].$$

Luego tomando $\{a_k\}_{k\in\mathbb{N}}$ tal que $\sum_{k=1}^\infty a_k c(k) < \infty$ se tiene que $\mu_s(\mathcal{A}^{-0})$ es finito. Para mostrar que μ_s es equivalente a μ notemos que de la definición de μ_s y CL se tiene

$$\int_{CL(E)}\int_0^\infty \mathbb{P}_\eta(X_s \in E)dsd\mu(\eta) \ = \ \int_{\mathcal{A}^{-0}}\int_0^\infty \mathbb{P}_\eta(X_s \in E)dsd\mu(\eta) \ = \mu_s(E),$$

pues de la Definición 4.1 se tiene

$$\eta \in CL(E) \iff \int_0^\infty \mathbb{P}_\eta(Y_t \in E)dt > 0,$$

y en particular $\mu_s(E) = 0 \iff \mu(CL(E)) = 0 \iff \mu(E) = 0$ y esta última equivalencia se tiene pues μ está bien concentrada. Veamos finalmente que μ_s cumple la ecuación (4.7), tomando para ello $A \subseteq \mathcal{A}^{-0}$ medible y notando que para $t \geq 0$

$$\int_{\mathcal{A}^{-0}} P_t(\eta, A)d\mu_s(\eta) \ = \ \int_{\mathcal{A}^{-0}}\int_0^\infty\int_{\mathcal{A}^{-0}} P_t(\eta, A)P_s(\eta', d\eta)dsd\mu(\eta'),$$

luego gracias a la propiedad de semigrupo se tiene

$$\begin{aligned}
\int_{\mathcal{A}^{-0}} P_t(\eta, A)d\mu_s(\eta) \ &= \ \int_{\mathcal{A}^{-0}}\int_0^\infty P_{t+s}(\eta', A)dsd\mu(\eta') \\
&\leq \ \int_{\mathcal{A}^{-0}}\int_0^\infty P_s(\eta', A)dsd\mu(\eta') \\
&= \ \mu_s(A).
\end{aligned}$$

Se concluye que μ_s cumple (4.7) de donde se sigue el resultado.

□

4.3. R-clasificación de procesos a tiempo discreto

En esta sección seguiremos el trabajo desarrollado por Tweedie en [27] y [28] sobre la R-clasificación de cadenas de Markov absorbidas a tiempo discreto, el cual se basa en las nociones de Harris-recurrencia y los trabajos de Harris en [12] y Vere-Jones en [31] y [32]. En este contexto trabajaremos con un proceso de Markov absorbido $\{X_n\}_{n\in\mathbb{N}}$ a valores sobre $\mathcal{X}$ y con semigrupo $\{P_n\}_{n\in\mathbb{N}}$. Asumiremos que X es φ-irreducible para una medida de probabilidad φ maximal.

En el Teorema 3.3 se estudió la equivalencia entre las distribuciones cuasi-estacionarias y los vectores propios por la izquierda del kernel de transición P_1, de ahí el interés por los vectores propios tanto por la derecha como por la izquierda de este operador. Con el fin de buscar la existencia de tales vectores propios, estudiaremos una clase más grande de objetos.

Definición 4.10 *Sea $r > 0$ y Q una medida σ-finita no nula sobre $\mathcal{X}$, diremos que Q es r-subinvariante si para todo A medible*

$$r\int_{\mathcal{X}} P(x,A)dQ(x) \ \leq \ Q(A),$$

y diremos que es r-invariante si la desigualdad se cumple φ-c.s. como igualdad. Así mismo, dada una función medible $f \geq 0$ no nula sobre un conjunto de medida φ positiva, diremos que es r-subinvariante si φ-casi seguramente

$$r\int_{\mathcal{X}} f(y)P(x,dy) \ \leq \ f(x),$$

y de igual forma, diremos que es r-invariante si la igualdad se alcanza φ-c.s. Si las desigualdades son en el sentido contrario, diremos que f y Q son r-superinvariantes.

El estudio de las medidas r-subinvariantes está fuertemente ligado al análisis de la serie de potencias $G_r : \mathcal{X} + \times \mathcal{B}_{\mathcal{X}} \to \mathbb{R}$ para $r \geq 0$ en que

$$G_r(x,A) \ := \ \sum_{n=1}^{\infty} r^n P^{(n)}(x,A).$$

Esta serie es análoga a la estudiada por Vere-Jones en [31] para el caso $\mathcal{X} = \mathbb{Z}$, y el siguiente resultado muestra que bajo la hipótesis de φ-irreducibilidad X posee una propiedad de solidaridad similar a la enunciada en ese trabajo.

Teorema 4.3 *[27, Teorema 1] Si X es φ-irreducible con φ maximal, entonces existe $R > 1$, una partición numerable $\mathcal{K}$ de $\mathcal{X}$ y $N \subseteq \mathcal{X}$ con $\varphi(N) = 0$ tal que*

1. *para todo $x \notin N$, $K \in \mathcal{K}$ y $A \subseteq K$ medible con $\varphi(A) > 0$, el radio de convergencia de la serie $G_{(\cdot)}(x,A)$ es R;*
2. *si existen $x \in \mathcal{X}$ y $B \subseteq \mathcal{X}$ medible tales que $G_R(x,B) < \infty$ entonces para todo $A \subseteq \mathcal{X}$ como antes y $x \notin N$ se tiene $G_R(x,A) < \infty$.*

Llamaremos a R el *radio de convergencia* del proceso, el cual está bien definido φ-c.s. gracias al teorema anterior. Notemos que la serie $G_1(x, A)$ corresponde al tiempo promedio que pasa el proceso en el conjunto A al comenzar en x, luego, como el proceso es absorbido, se debe tener que $G_1(x, A) < \infty$.

Observación Gracias al teorema anterior, si existen $r > 0$, A medible y $x \in \mathcal{X}$ tales que $G_r(x, A) < \infty$, entonces la función

$$G_r(\cdot, A) : \mathcal{X} \longrightarrow \mathbb{R}$$

está bien definida φ-c.s. y es r-subinvariante. Así mismo, la medida

$$G_r(x, \cdot) : \mathcal{B}_{\mathcal{X}} \longrightarrow \mathbb{R}$$

es σ-finita y r-subinvariante.

De la observación anterior y el Teorema 4.3, para $r < R$ siempre existen funciones y medidas r-subinvariantes. El siguiente resultado muestra que la conexión entre R y tales medidas y funciones es aún más fuerte.

Proposición 4.9 *[27, Proposición 3.1] Sea $r > R$, entonces no existen medidas ni funciones r-subinvariantes.*

El caso $r = R$ es mucho más delicado y para poder estudiar este caso es necesario introducir las siguientes definiciones.

Definición 4.11 *Sea R el radio de convergencia de X. Diremos que tal proceso es*

1. *R-transiente si existen $x \in \mathcal{X}$ y A medibles con $G_R(x, A) < \infty$;*

2. *R-recurrente si para todo $x \in \mathcal{X}$ y A medible se cumple $G_R(x, A) = \infty$.*

Gracias al Teorema 4.3, X debe ser R-recurrente o R-transiente, y en este último caso gracias a la observación anterior G_R define funciones y medidas R-subinvariantes. Uno de los principales resultados obtenidos por Tweedie en [27] muestra la existencia de tales funciones y medidas en el caso R-recurrente.

Teorema 4.4 *[27, Proposición 3.4 y teorema 4] Si X es R-recurrente, entonces*

1. *existe una única función (salvo ponderación) $f : \mathcal{X} \to \mathbb{R}$ medible Λ^*-subinvariante. Tal función es Λ^*-invariante;*

2. *existe una única medida (salvo ponderación) $Q : \mathcal{B}_{\mathcal{X}} \to \mathbb{R}$ σ-finita Λ^*-subinvariante. Tal medida es Λ^*-invariante y equivalente a φ.*

A f y Q como en el teorema los llamaremos *vectores propios maximales*, cuya existencia está garantizada para el caso R-recurrente. Reciprocamente, veremos que bajo una hipótesis de integrabilidad, la existencia de vectores propios maximales garantiza la R-recurrencia del proceso.

Proposición 4.10 *[28, Proposición 10.4] Sea f una función medible r-superinvariante y Q una medida σ-finita r-subinvariante. Si se cumple*

$$\int_{\mathcal{X}} f(x)dQ(x) \ < \ \infty,$$

entonces $R = r$ y el proceso es R-recurrente, luego gracias al Teorema 4.4, Q es R-invariante y única salvo ponderación.

La condición de integrabilidad de los vectores propios maximales es suficiente pero no necesaria para que el proceso sea R-recurrente, y en general existen algunas diferencias interesantes entre aquellos procesos R-recurrentes que la cumplen y aquellos que no. Esto motiva la siguiente subclasificación.

Definición 4.12 *Supongamos que X es R-recurrente con vectores propios maximales f y Q. Diremos que el proceso es*

1. *R-recurrente nulo si $\int_{\mathcal{X}} f(x)dQ(x) = \infty$;*
2. *R-recurrente positivo si $\int_{\mathcal{X}} f(x)dQ(x) < \infty$.*

El siguiente teorema ilustra la diferencia entre ambos casos.

Teorema 4.5 *[27, Teorema 6] Supongamos que X es R-recurrente y aperiódico (ver [25]), con vectores propios maximales f y Q. Entonces existe una partición $\mathcal{K}$ tal que para todo $K \in \mathcal{K}$ y todo $A \subseteq K$ medible con $\varphi(A) > 0$ se cumple*

$$R^n P^{(n)}(x, A) \xrightarrow[n\to\infty]{} \frac{f(x)Q(A)}{\int_{\mathcal{X}} f(y)dQ(y)} \qquad \varphi\text{-c.s. en } \mathcal{X},$$

en que el término de la derecha es nulo si el proceso es R-recurrente nulo.

De esta forma, para el caso R-positivo es posible hablar de un proceso *asintótico*, el cual corresponde al proceso condicionado a la no extinción.

Las medidas r-subinvariantes estudiadas en esta sección son σ-finitas, y por lo tanto tales medidas no corresponden necesariamente a distribuciones cuasi-estacionarias pues estas últimas son de probabilidad. Nos interesa entonces saber para qué valores de $r \leq R$ existen medidas de probabilidad r-subinvariantes, lo que motiva la introducción del parámetro $R_1 \leq R$, definido en [2] que adaptamos para procesos a tiempo discreto como

$$R_1 \ := \ \sup\{r \geq 1, \ \exists x \in \mathcal{X}, \ G_r(x, \mathcal{X}) < \infty\}.$$

El siguiente resultado sigue de adaptar lo hecho en [2] a procesos a tiempo discreto, y permite ilustrar la conexión entre R_1 y las q.s.d. de la siguiente forma.

Lema 4.1 *[2, Lema 3.5] Para R_1 se tienen las siguientes caracterizaciones;*

- $R_1 = \sup\{r \geq 1, \ \textit{existe una medida de probabilidad } r\textit{-subinvariante}\}$;
- $R_1 = \sup\{r \geq 1, \ \textit{existe una función } f \ r\textit{-subinvariante con } 1 \leq f < \infty\}$.

Nos interesa particularmente el caso en que X es R_1-recurrente, pues en tal caso existe la posibilidad de que la única medida R_1-recurrente sea una medida de probabilidad.

4.4. Λ^*-clasificación de procesos a tiempo continuo

Los resultados que presentamos a continuación siguen muy de cerca lo hecho para procesos a tiempo discreto, y la totalidad de ellos corresponden a teoremas demostrados o adaptados por Breyer en [2]. Trabajaremos con un proceso X absorbido a tiempo continuo sobre un espacio $\mathcal{X}$ tal que existe una medida de irreducibilidad maximal φ.

Definición 4.13 *Sea $\theta > 0$ y ν una medida σ-finita no nula sobre $\mathcal{X}$, diremos que ν es θ-subinvariante si para todo $0 \leq f < \infty$ y todo $t \geq 0$ se cumple*

$$\langle \nu, P_t(f) \rangle \ \leq \ e^{-\theta t} \langle \nu, f \rangle,$$

y diremos que es θ-invariante si la desigualdad se cumple como igualdad para toda f. De igual forma, diremos que una función medible no negativa, f es θ-subinvariante si φ-casi seguramente sobre $\mathcal{X}$ y para todo $t \geq 0$

$$P_t(f) \ \leq \ e^{-\theta t} f,$$

y diremos que es θ-invariante si la igualdad se alcanza φ-c.s. Si las desigualdades son en el sentido contrario, diremos que la medida y la función anteriores son θ-superinvariantes.

Estas definiciones son análogas a las hechas en la sección anterior, y no es extraño que al definir el kernel $V_\lambda : \mathcal{X} \times \mathcal{B}_\mathcal{X} :\to \mathbb{R}$ de forma análoga a G_r como

$$V_\lambda(x, A) \ := \ \int_0^\infty e^{\lambda t} P_t(x, A) dt,$$

se tenga el siguiente teorema de solidaridad:

Teorema 4.6 *[2, Teorema 3.3] Existe $\Lambda^* \geq 0$, tal que para todo $\lambda > \Lambda^*$, $x \in \mathcal{X}$ y $A \subseteq \mathcal{X}$,*

$$V_\lambda(x, A) = \infty,$$

y existen $N \subseteq \mathcal{X}$ φ-nulo y una partición numerable $\mathcal{K}$ de $\mathcal{X}$ tales que para todo $\lambda < \Lambda^*$, $x \notin N$, $K \in \mathcal{K}$ y $A \subseteq K$ con medida positiva,

$$V_\lambda(x, A) < \infty.$$

Si existen $x \in \mathcal{X}$ y $A \subseteq \mathcal{X}$ tales que $V_\lambda(x, A) < \infty$, diremos que V_λ es transiente, y en caso contrario, que es recurrente. El proceso X se clasifica de acuerdo a lo que ocurre con V_{Λ^*}, es decir, si V_{Λ^*} es transiente X es Λ^*-transiente, y en caso contrario X es Λ^*-recurrente. El siguiente resultado es una adaptación del Teorema 4.4 para procesos a tiempo continuo.

Teorema 4.7 *[2, Teorema 3.10] Supongamos que X es Λ^*-recurrente, entonces*

1. *existe una única función (salvo ponderación) $0 \leq f < \infty$ medible Λ^*-subinvariante. Tal función es Λ^*-invariante;*
2. *existe una única medida σ-finita Λ^*-subinvariante ν. Tal medida es Λ^*-invariante y equivalente a φ.*

De igual forma, la siguiente es una adaptación de la Proposición 4.10.

Proposición 4.11 *[2, Proposición 3.11] Sea $0 \leq f < \infty$ una función medible λ-superinvariante y ν una medida σ-finita λ-subinvariante. Si se cumple*

$$\langle \nu, f \rangle \ < \ \infty,$$

entonces $\Lambda^* = \lambda$ y el proceso es Λ^*-recurrente positivo.

Tal y como en la sección anterior llamamos a f y ν los *vectores propios maximales* del proceso, lo que permite clasificar el proceso como Λ^*-recurrente positivo o Λ^*-recurrente nulo de forma análoga a lo hecho para procesos a tiempo discreto. Así mismo se define Λ_1 como

$$\Lambda_1 \ := \ \sup\{\lambda \geq 0, \ \exists x \in \mathcal{X}, \ V_\lambda(x, \mathcal{X}) < \infty\},$$

para el cual se tiene la versión a tiempo continuo del Lema 4.1.

Lema 4.2 *[2, Lema 3.5] Para Λ_1 definido como antes se tienen las siguientes caracterizaciones;*

- $\Lambda_1 = \sup\{\lambda \geq 0,$ *existe una medida de probabilidad λ-subinvariante*$\}$;
- $\Lambda_1 = \sup\{\lambda \geq 0,$ *existe una función f λ-subinvariante con $1 \leq f < \infty$*$\}$.

4.5. Resultados para el caso φ-s.s.i.

En la presente sección mostraremos que la propiedad de ser φ-s.s.i. permite clasificar un proceso absorbido a tiempo continuo mirando únicamente sus esqueletos a tiempo discreto, y que además esta propiedad proporciona condiciones equivalentes a la Λ^*-positividad que no involucran el parámetro Λ^*. La totalidad de los resultados aquí enunciados son desarrollados por Tweedie en [26].

Al igual que en la sección anterior tomamos trabajamos con un proceso absorbido a tiempo continuo X sobre $\mathcal{X}$, y asumiremos además que X es φ-s.s.i. con φ maximal. Utilizaremos la notación P^h para referirnos al semigrupo $\{P_{nh}\}_{n\in\mathbb{N}}$.

Teorema 4.8 *[26, Teorema 4] Si X es φ-s.s.i. con φ maximal, entonces para $\lambda \geq 0$,*

1. *V_λ es λ-recurrente (positivo) si y sólo si para todo $h > 0$, P^h es $e^{\lambda h}$-recurrente (positivo);*
2. *Si f es $e^{h\lambda}$-subinvariante para P^h, entonces f es λ-invariante para X. Similarmente, si ν es $e^{h\lambda}$-subinvariante para P^h, entonces ν es λ-invariante para X.*

Utilizando este teorema, deducimos que la Λ^*-clasificación de $\{X_t\}_{t\geq 0}$ es igual a la R-clasificación de $\{X_n\}_{n\in\mathbb{N}}$ y que $R = e^{\Lambda^*}$.

Asumiendo que el proceso X es absorbido casi-seguramente se definen;

$$\begin{aligned}
a(x,A,t) &:= \frac{1}{P_t(\gamma)(x)}\int_A \gamma(y)P_t(x,dy) \;=\; \mathbb{P}_x(X_t \in A \,\big|\, t < \mathcal{T} < \infty);\\
s(x,A,t,u) &:= \frac{1}{P_{t+u}(\gamma)(x)}\int_A P_u(\gamma)(y)P_t(x,dy) \;=\; \mathbb{P}_x(X_t \in A \,\big|\, t+u < \mathcal{T} < \infty);\\
b(x,y,t) &:= \frac{P_t(\gamma)(x)}{P_t(\gamma)(y)} \;=\; \frac{\mathbb{P}_x(t < \mathcal{T} < \infty)}{\mathbb{P}_y(t < \mathcal{T} < \infty)},
\end{aligned}$$

en que $x, y \in \mathcal{X}$, $A \in \mathcal{B}_{\mathcal{X}}$ y $t, u > 0$. Utilizando estas funciones podemos encontrar una condición equivalente que el proceso sea Λ^*-recurrente positivo sin la necesidad de conocer el parámetro Λ^*.

Teorema 4.9 *[26, Teorema 7] Supongamos que el proceso X es φ-s.s.i. entonces son equivalentes:*

- *existe un conjunto N φ-nulo tal que para todo $x, y \notin N$, A medible y $u > 0$ los límites*

$$\lim_{t\to 0} a(x,A,t) \quad y \quad \lim_{t\to\infty} s(x,A,t,u)$$

  existen y definen medidas de probabilidad en $\mathcal{X}$, y el límite $\lim_{t\to 0} b(x,y,t)$ existe y no es cero φ-c.s.;

- *X es Λ^*-positivo y existe una única medida de probabilidad Λ^*-invariante.*

En el caso en que el proceso X es Λ^*-positivo con vectores propios maximales π y f en que π es de probabilidad, podemos expresar a partir de éstos los límites anteriores como

$$\begin{aligned}
\lim_{t\to\infty} a(x,A,t) &= \pi(A),\\
\lim_{t\to\infty}\lim_{u\to\infty} s(x,A,t,u) &= \frac{1}{\langle \pi, f\rangle}\int_A f(z)d\pi(z),\\
\lim_{t\to\infty} b(x,y,t) &= \frac{f(x)}{f(y)}.
\end{aligned}$$

Capítulo 5

El proceso truncado

Con el fin de modelar de forma sencilla la evolución de poblaciones, usualmente se asume que el tamaño de la misma puede ser arbitrariamente grande, sin embargo en muchos casos este supuesto es un absurdo pues el sistema en el que se encuentran tales individuos tiene una capacidad limitada. Al llegar al límite de esta capacidad puede ocurrir que el sistema colapsa, siendo incapaz de sustentar a la población y por lo tanto ésta se extingue, o que los individuos dejan de reproducirse debido a la falta de recursos. Considerando estos casos, no es ilógico modelar la evolución de poblaciones de forma tal que la cantidad de individuos se encuentre acotada, independientemente de los rasgos presentes.

En este capítulo estudiamos el proceso $\hat{Y}$, similar a Y salvo en que el espacio de estados se ve reducido a $\mathcal{A}_{\leq p}$ para algún $p \in \mathbb{N}$. Matemáticamente la ventaja de considerar este caso particular es que el espacio de estados pasa a ser compacto, luego tanto la cantidad de individuos como las tasas de transición se encuentran acotadas, lo que proporciona una condición ideal para aplicar los resultados obtenidos en los capítulos anteriores. Comenzamos este estudio definiendo el proceso truncado y observando que la mayoría de los resultados obtenidos en el primer capítulo se mantienen válidos, luego introducimos una familia de procesos a la que llamamos *aproximaciones simples* y la cual utilizamos para encontrar un criterio suficiente que garantice la unicidad de la distribución cuasi-estacionaria para $\hat{Y}$.

5.1. Construcción de $\hat{Y}$

En la definición 2.5 se definieron las tasas de transición como funciones que toman valores estrictamente positivos sobre $\mathbb{T} \times \mathcal{A}^{-0}$ y valores nulos fuera de este conjunto, suposición que debilitamos ligeramente al permitir la posibilidad de que la función $b + m$ alcance valores nulos sobre el conjunto $\mathbb{T} \times \mathcal{A}_p$, y asumiendo que λ es nulo sobre $\mathbb{T} \times \mathcal{A}_{p+1}$. Construimos entonces el proceso $\hat{Y}$ a partir de Y de la siguiente manera.

Definición 5.1 *Dado $\mathcal{T}_{p+1}$ como en la Definición 2.19 se define $\hat{Y}$ como*

$$\hat{Y}_t := Y_{t \wedge \mathcal{T}_{p+1}}.$$

Observación $\hat{Y}$ es un proceso de saltos con la propiedad de Markov, el cual es absorbido casi-seguramente en $\{0\} \cup \mathcal{A}_{>p}$. Si suponemos además que $Y_0 \in \mathcal{A}_{\leq p}$, entonces en $[0, \mathcal{T}^Y_{p+1}]$ se tiene $\hat{Y}_t = Y_t$, luego utilizando la notación introducida en la Definición 2.19 es directo que

- $\mathcal{T}^{\hat{Y}}_{p+1} = \mathcal{T}^Y_{p+1}$;
- $\tau^{\hat{Y}} = \tau^Y$;
- $N^{\hat{Y}}_t \leq N^Y_t$,

y si $Y_0 \notin \mathcal{A}_{\leq p}$, entonces $T^{\hat{Y}}_{p+1} = 0$, $\tau^{\hat{Y}} = \infty$ y $N^{\hat{Y}}_t = 0$.

Utilizando la observación anterior, en el caso en que $Y_0 \in \mathcal{A}_{\leq p}$ utilizaremos τ y $\mathcal{T}_n$ para referirnos al tiempo del primer salto y de llegada a $\mathcal{A}_n$ tanto para Y como para $\hat{Y}$. Notando que si Y comienza en $\mathcal{A}_{\leq p}$, necesariamente se tiene $\tau \leq \mathcal{T}_{p+1}$ y por lo tanto $\hat{Y}_\tau = Y_\tau$, es directo que el kernel de transición $\hat{Q}$ del proceso $\hat{Y}$ debe cumplir $\hat{Q}(Y_0, \cdot) \equiv Q(Y_0, \cdot)$.

Por otro lado, si $Y_0 \notin \mathcal{A}_{\leq p}$ entonces $\tau^{\hat{Y}} = \infty$, de donde necesariamente $\hat{Q}(Y_0, \cdot) \equiv 0$, y entonces podemos pensar $\hat{Q}$ como el kernel de transición de un proceso construido como en la Definición 2.11 utilizando para ello las tasas $b\mathbf{1}_{\mathcal{A}_{\leq p}}$, $m\mathbf{1}_{\mathcal{A}_{\leq p}}$ y $\lambda\mathbf{1}_{\mathcal{A}_{\leq p}}$.

La flexibilidad que hemos otorgado a la función $b + m$ sobre el conjunto $\mathbb{T} \times \mathcal{A}_p$ permite desarrollar la teoría sin la necesidad de especificar cómo se trunca Y. Si $(b + m)|_{\mathcal{A}_p} \equiv 0$, se habla de un proceso con una barrera reflectante en el nivel p, es decir, si se llega a $\mathcal{A}_p$, en el siguiente salto el proceso vuelve a $\mathcal{A}_{p-1}$. Por otro lado, si $b + m > 0$ en $\mathcal{A}_p$ se habla de un proceso que se extingue en $\mathcal{A}_{p+1}$, el cual representa el caso en que colapsa el medioambiente debido a la sobrepoblación. La posibilidad de que $b + m$ que tenga tanto valores nulos como positivos permite considerar casos en que el comportamiento de la población frente a la sobrepoblación depende de la configuración y los rasgos de los individuos que la componen.

A continuación mostraremos brevemente que las propiedades obtenidas en el segundo capítulo para Y siguen siendo válidas para $\hat{Y}$. Para ello comenzamos observando que $\hat{Y}$ no posee explosiones.

Proposición 5.1 *Para todo $t \geq 0$ se cumple* $\mathbb{P}\left(N^{\hat{Y}}_t = \infty\right) = 0$.

Demostración. Utilizando la Proposición 2.7 se tiene que $\mathbb{P}(N^Y_t = \infty) = 0$, y por otro lado, de la obseración anterior se tiene que $N^{\hat{Y}}_t \leq N^Y_t$, de donde se concluye

$$\mathbb{P}\left(N^{\hat{Y}}_t = \infty\right) \leq \mathbb{P}\left(N^Y_t = \infty\right) = 0.$$

□

Para la absorción exponencial de $\hat{Y}$ la demostración es tan evidente como la anterior. En efecto, definiendo $\mathcal{T}^{\hat{Y}}$ el tiempo de absorción de $\hat{Y}$ en su cementerio se tiene

$$\mathcal{T}^{\hat{Y}} = \mathcal{T}^{\hat{Y}}_0 \wedge \mathcal{T}^{\hat{Y}}_{p+1} \leq \mathcal{T}^Y_{p+1} \wedge \mathcal{T}^Y_0 \leq \mathcal{T}^Y_0,$$

luego el siguiente resultado es evidente.

Proposición 5.2 *Existen θ_0 y $c > 0$ tales que para todo $\theta < \theta_0$ y todo $\eta \in \mathcal{A}_{\leq p}$*

$$\mathbb{E}_\eta\left(e^{\theta\mathcal{T}^{\hat{Y}}}\right) \leq c.$$

Ya hemos visto que $\hat{Y}$ puede pensarse como el proceso Y pero en que las tasas de transición se anulan para configuraciones de más de p individuos. Así, lo lógico sería definir débilmente el generador infinitesimal $\hat{G}$ de $\hat{Y}$ como

$$\begin{aligned}\hat{G}f(\eta) \;&:=\; \mathbf{1}_{\mathcal{A}_{\leq p}}(\eta)\left(\sum_{y\in\{\eta\}}\eta_y b_y(\eta)[f(\eta+\delta_y)-f(\eta)] + \sum_{y\in\{\eta\}}\eta_y\lambda_y(\eta)[f(\eta-\delta_y)-f(\eta)]\right.\\ &\qquad\left. + \sum_{y\in\{\eta\}}\eta_y m_y(\eta)\int_{\mathbb{T}}[f(\eta+\delta_z)-f(\eta)]g(y,z)d\sigma(z)\right),\end{aligned}$$

para toda función f que se anule fuera de $\mathcal{A}_{\leq p}$. Mostraremos a continuación que en efecto se tiene este resultado.

Proposición 5.3 *Sea $f : \mathcal{A} \to \mathbb{R}$ acotada tal que se anula fuera de $\mathcal{A}_{\leq p}$, entonces el proceso*

$$\mathcal{M}_t^p \;:=\; f(\hat{Y}_t) - f(\hat{Y}_0) - \int_0^t \hat{G}f(\hat{Y}_s)ds,$$

es una martingala.

Demostración. Como la función f está acotada, podemos utilizar la Proposición 2.10 para deducir que el proceso

$$\mathcal{M}_{t\wedge\mathcal{T}_{p+1}} \;:=\; f(Y_{t\wedge\mathcal{T}_{p+1}}) - f(Y_0) - \int_0^{t\wedge\mathcal{T}_{p+1}} Gf(Y_s)ds,$$

es una martingala. Ahora, por definición de $\hat{Y}$ esta expresión es igual a

$$f(\hat{Y}_t) - f(\hat{Y}_0) - \int_0^{t\wedge\mathcal{T}_{p+1}} \mathbf{1}_{\mathcal{A}_{\leq p}}(\hat{Y}_s)Gf(\hat{Y}_s)ds,$$

donde hemos utilizado que para $s < \mathcal{T}_{p+1}$ se tiene $Y_s = \hat{Y}_s \in \mathcal{A}_{\leq p}$. Por otro lado, para $s \geq \mathcal{T}_{p+1}$ se tiene $\mathbf{1}_{\mathcal{A}_{\leq p}}(\hat{Y}_s) = 0$, luego

$$\mathcal{M}_{t\wedge\mathcal{T}_{p+1}} \;=\; f(\hat{Y}_t) - f(\hat{Y}_0) - \int_0^t \mathbf{1}_{\mathcal{A}_{\leq p}}(\hat{Y}_s)Gf(\hat{Y}_s)ds \;=\; \mathcal{M}_t^p,$$

y se concluye que $\mathcal{M}_t^p$ es una martingala.

□

Una vez que el proceso $\hat{Y}$ llega al conjunto cementerio se comporta como una constante y además , en virtud del Teorema 3.2, el estudio de las q.s.d. para este proceso es independiente de dónde fue absorbido, lo que motiva colapsar el cementerio en el singleton $\{0\}$. Para ello es necesario adaptar los elementos asociados a $\hat{Y}$ de la siguiente manera.

Definición 5.2 *Para el proceso $\hat{Y}$ utilizaremos las siguientes notaciones:*

1. *dado $\eta \in \mathcal{A}_{\leq p}$ utilizamos la notación $\hat{Q}(\eta, 0)$ para referirnos a $Q(\eta, \{0\} \cup \mathcal{A}_{p+1})$;*
2. *dada $f : \mathcal{A}_{\leq p} \to \mathbb{R}$ se define $\bar{f} : \mathcal{A} \to \mathbb{R}$ la extensión de f que es nula fuera de $\mathcal{A}_{\leq p}$;*
3. *dada $f : \mathcal{A}_{\leq p} \to \mathbb{R}$ se define $\hat{G}(f)$ como*
$$\hat{G}(f) \ := \ G(\bar{f})\big|_{\mathcal{A}_{\leq p}};$$
4. *dados $\eta \in \mathcal{A}_{\leq p}$ y $A \subseteq \mathcal{A}_{\leq p}$ se define el semigrupo de transición de $\hat{Y}$, $\{\hat{P}_t\}_{t\geq 0}$ como*
$$\hat{P}_t(\eta, A) \ := \ \mathbb{P}_\eta(Y_t \in A, \ \mathcal{T}_{p+1} > t) \ = \ \mathbb{P}_\eta(Y_t \in A),$$
en que la igualdad se tiene pues λ es nulo en $\mathbb{T} \times \mathcal{A}_{p+1}$, y al igual que para $\hat{G}$ utilizamos la notación $\hat{P}_t(f) := \hat{P}_t(\bar{f})\big|_{\mathcal{A}_{\leq p}}$.

Notemos que como $\mathcal{A}_{\leq p}$ es un abierto y un cerrado, dada $f \in \mathcal{C}(\mathcal{A}_{\leq p})$ se tiene $\bar{f} \in \mathcal{C}_b(\mathcal{A})$ y por lo tanto, gracias a la Proposición 3.5 se tiene el siguiente resultado.

Proposición 5.4 *Para toda función $f \in \mathcal{C}(\mathcal{A}_{\leq p})$ y todo $t \geq 0$ se tiene $\hat{P}_t(f) \in \mathcal{C}(\mathcal{A}_{\leq p})$.*

Demostración. Ya habíamos visto en la Definición 5.2 que para $\eta \in \mathcal{A}_{\leq p}$ y $A \subseteq \mathcal{A}_{\leq p}$ se tiene $\hat{P}(\eta, A) = P(\eta, A)$ de donde se deduce

$$\hat{P}_t\bar{f}(\eta) \ = \ \int_{\mathcal{A}_{\leq p}} f(\eta')d\hat{P}_t(\eta, \eta') \ = \ \int_{\mathcal{A}_{\leq p}} f(\eta')dP_t(\eta, \eta') \ = \ P_t\bar{f}(\eta),$$

y como $\bar{f}$ es continua y acotada, utilizando la Proposición 3.5 se deduce que $\hat{P}_t\bar{f}$ es continua y acotada luego se concluye el resultado.

□

Como $\hat{Y}$ es un proceso que posee la propiedad de Feller y que toma valores sobre un espacio métrico compacto, podemos utilizar la Proposición 3.4 para concluir la existencia de distribuciones cuasi-estacionarias para $\hat{Y}$.

Proposición 5.5 *Existe al menos una distribución cuasi-estacionaria para $\hat{Y}$, la cual se concentra en $\mathcal{A}_{\leq p}$.*

Hemos probado así que la totalidad de los resultados encontrados para Y en los capítulos 2 y 3 se mantienen al truncar el proceso. Veremos a continuación que esto también es cierto para los resultados encontrados en el capítulo anterior.

Proposición 5.6 *Dado $B \subseteq \mathcal{A}_{\leq p}$ medible, los siguientes conjuntos son iguales;*

1. $\left\{\eta \in \mathcal{A}_{\leq p},\ \exists n \in \mathbb{N},\ \mathbb{P}_\eta(\bar{Y}_n \in B,\ \mathcal{T}_{p+1} > \tau_n) > 0\right\}$;
2. $\left\{\eta \in \mathcal{A}_{\leq p},\ \forall t > 0,\ \hat{P}_t(\eta, B) > 0\right\}$;
3. $\left\{\eta \in \mathcal{A}_{\leq p},\ \exists t > 0,\ \hat{P}_t(\eta, B) > 0\right\}$;
4. $CL_{\hat{Y}}(B)$.

Demostración. Ya habíamos visto en la Definición 5.2 que para $\eta \in \mathcal{A}_{\leq p}$ y $B \subseteq \mathcal{A}_{\leq p}$ se tiene $\hat{P}_t(\eta, B) = P_t(\eta, B)$, y como λ es nulo en $\mathbb{T} \times \mathcal{A}_{p+1}$, para todo $\eta \notin \mathcal{A}_{\leq p}$ se tiene $P_t(\eta, B) = 0$. De esta forma deducimos

$$\hat{P}_t(\eta, B) > 0 \iff P_t(\eta, B) > 0 \ \wedge\ \eta \in \mathcal{A}_{\leq p},$$

y equivalentemente

$$\mathbb{P}_\eta(\bar{Y}_n \in B,\ \mathcal{T}_{p+1} > \tau_n) > 0 \iff \mathbb{P}_\eta(\bar{Y}_n \in B) > 0 \ \wedge\ \eta \in \mathcal{A}_{\leq p}.$$

Es sencillo ver entonces que los conjuntos 1, 2, 3 y 4 son iguales a los respectivos conjuntos definidos en la Proposición 4.2, los que son iguales entre sí gracias a la misma proposición. □

Concluiremos esta sección enunciando el siguiente resultado, según el cual la medida μ se mantiene como la medida de irreducibilidad maximal de $\hat{Y}$, y cuya demostración omitiremos pues es idéntica a lo hecho en el Teorema 4.1.

Proposición 5.7 *Sea $E \subseteq \mathcal{A}_{\leq p}$ medible, entonces:*

1. *si $\mu(E) = 0$, entonces $\mu(CL_{\hat{Y}}(E)) = 0$;*
2. *si $\mu(E) > 0$, entonces $CL_{\hat{Y}}(E) = \mathcal{A}_{\leq p}$.*

La totalidad de las propiedades obtenidas para el proceso Y en el capítulo anterior se basan en la Proposición 4.2 y el Teorema 4.1, luego las Proposiciones 5.6 y 5.7 permiten concluir que los mismos resultados pueden aplicarse a $\hat{Y}$. Nos interesa especialmente que $\hat{Y}$ es μ-s.s.i. y que podemos aplicar el teorema 4.2 al semigrupo $\{\hat{P}_t\}_{t\geq 0}$.

5.2. Aproximaciones para $\hat{Y}$

En la sección anterior, gracias a la Proposición 5.5 existe una distribución cuasi-estacionaria para el proceso $\hat{Y}$, sin embargo de la demostración utilizada para su existencia no podemos desprender nada respecto a las propiedades de tal distribución y por lo tanto es necesario un enfoque distinto para su estudio. El enfoque utilizado en este capítulo es definir una familia de procesos $\{\hat{W}^n\}_{n\in\mathbb{N}}$ que aproximen a $\hat{Y}$, y cuyas propiedades podamos extender a este proceso.

La idea de la construcción de cada $\hat{W}^n$ consiste en definir una partición finita de *clases de equivalencia* de rasgos, la cual induce una partición numerable en $\mathcal{A}^{-0}$, y que iremos tomando cada vez más finas. Definiremos entonces un proceso $\hat{W}^n$ cuyas tasas sean constantes en cada clase de equivalencia y que por lo tanto se comporte como un proceso a espacio de estados discretos.

Definición 5.3 *Se define la familia $\{\mathcal{D}_n\}_{n\in\mathbb{N}}$ de particiones finitas de $\mathbb{T}$, la cual cumple:*

1. *para todo $n \in \mathbb{N}$ y $D \in \mathcal{D}_n$ se tiene $\sigma(D) > 0$;*
2. *para todo $n \in \mathbb{N}$, $D \in \mathcal{D}_n$ y todo par $y_1, y_2 \in D$ se cumple $d_{\mathbb{T}}(y_1, y_2) \leq \frac{1}{n}$.*

La existencia de $\{\mathcal{D}_n\}_{n\in\mathbb{N}}$ es directa de la compacidad de $\mathbb{T}$ y de que el soporte de σ coincide con este conjunto. Veremos a continuación que $\{\mathcal{D}_n\}_{n\in\mathbb{N}}$ define una familia de particiones numerables sobre $\mathcal{A}^{-0}$.

Definición 5.4 *Dada $\{\mathcal{D}_n\}_{n\in\mathbb{N}}$ como en la definición anterior se define la familia $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ de particiones numerables y medibles de $\mathcal{A}^{-0}$ en que para todo $n \in \mathbb{N}$ y $H \in \mathcal{H}_n$ existen $k \in \mathbb{N}$ y $D_1, D_2, \ldots, D_k \in \mathcal{D}_n$ no necesariamente distintos con*

$$H = \left\{ \sum_{j=1}^{k} \delta_{z_j},\ z_1 \in D_1, z_2 \in D_2, \ldots, z_k \in D_k \right\}.$$

De esta manera, cada elemento de $\mathcal{H}_n$ es definido por una única secuencia finita de elementos en $\mathcal{D}_n$, cuyo largo corresponde a la cantidad de individuos en las configuraciones de H, y como cada D_i tiene medida positiva se tiene $\mu(H) > 0$.

Con el fin de facilitar la notación, definiremos para cada $n \in \mathbb{N}$ un orden en $\mathbb{T}$ inducido por $\mathcal{D}_n$.

Definición 5.5 *Sea $n \in \mathbb{N}$ y $D_1, D_2, \ldots, D_k$ una enumeración de los elementos en $\mathcal{D}_n$. Se define la relación de orden $\leq_n$ como*

$$y \leq_n z \iff (\exists i \leq j,\ y \in D_i,\ z \in D_j) \ \vee\ (\exists i,\ y, z \in D_i \ \wedge\ y \leq z),$$

en que $\leq$ corresponde al orden definido en el segundo capítulo. En lo que resta de este trabajo, para $\eta \in \mathcal{A}^{-0}$ asumiremos siempre que $\vec{\eta}$ está ordenado según $\leq_n$.

El orden $\leq_n$ será de utilidad para reducir la notación, pues asumiendo que los vectores $\vec{\eta}$ están ordenados según esta relación, para cada par $\vec{\eta_1}, \vec{\eta_2}$ con $\eta_1, \eta_2 \in H \in \mathcal{H}_n$, y cada j, los rasgos j-ésimos pertenecen a la misma clase en $\mathcal{D}_n$.

Proposición 5.8 *Sean $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ y $\{\mathcal{D}_n\}_{n\in\mathbb{N}}$ como en las Definiciones 5.3 y 5.4. Para todo $\varepsilon > 0$ podemos tomar n suficientemente grande, de forma de que para todo $H \in \mathcal{H}_n$ y $\eta_1, \eta_2 \in H$ en que $\vec{\eta_1} = (x_1, \dots, x_k)$ y $\vec{\eta_2} = (y_1, \dots, y_k)$, y para todo $1 \leq j \leq k$, $z \in \mathbb{T}$ se cumple*

$$|b_{x_j}(\eta_1) - b_{y_j}(\eta_2)| + |m_{x_j}(\eta_1) - m_{y_j}(\eta_2)| + |\lambda_{x_j}(\eta_1) - \lambda_{y_j}(\eta_2)| + |g(x_j, z) - g(y_j, z)| \leq \varepsilon e^k$$

Demostración. Sabemos que $e^{-\|\cdot\|}b, e^{-\|\cdot\|}m$ y $e^{-\|\cdot\|}\lambda$ son continuas sobre $\mathbb{T} \times \mathcal{A}^{-0}$ y además, gracias a las hipótesis (H) se tiene

$$e^{-\|\cdot\|}(b + m + \lambda) \xrightarrow[\|\eta\|\to\infty]{} 0,$$

luego, para todo $p \in \mathbb{N}$ existe $\delta_1 > 0$ tal que si $(y_1, \eta_1), (y_2, \eta_2) \in \mathbb{T}\times\mathcal{A}_{\leq p}$ cumplen la condición $d_{\mathbb{T}}(y_1, y_2) + d(\eta_1, \eta_2) \leq \delta_1$, entonces $\|\eta_1\| = \|\eta_2\|$ y

$$|b_{y_1}(\eta_1) - b_{y_2}(\eta_2)| + |m_{y_1}(\eta_1) - m_{y_2}(\eta_2)| + |\lambda_{y_1}(\eta_1) - \lambda_{y_2}(\eta_2)| \leq e^{\|\eta_1\|}\varepsilon/2,$$

y esta cota es válida también si $\|\eta_1\|, \|\eta_2\| > p$. Por otro lado, como $g : \mathbb{T}^2 \to \mathbb{R}$ es uniformemente continua, existe $\delta_2 > 0$ tal que

$$d_{\mathbb{T}}(y_1, y_2) \leq \delta \implies |g(y_1, z) - g(y_2, z)| \;\leq\; \varepsilon/2 \;\leq\; e^{\|\eta_1\|}\varepsilon/2.$$

Gracias a estas cotas y como $d_{\mathbb{T}}(y_1, y_2) \leq \frac{1}{n}$ para $y_1, y_2 \in D \in \mathcal{D}_n$ con n grande, bastará probar que en tal caso, para $\eta_1, \eta_2 \in H \in \mathcal{H}_n$ con $k \leq p$ individuos, el valor de $d(\eta_1, \eta_2)$ es pequeño. Para ello, siguiendo las definiciones introducidas en el segundo capítulo y utilizando la definición de la métrica de Lévy-Prokhorov, la distancia entre η_1 y η_2 está definida como

$$d(\eta_1, \eta_2) \;=\; k\,\text{ínf}\{\varepsilon > 0,\ \eta_1(A) \leq \eta_2(A^\varepsilon) + \varepsilon \;\wedge\; \eta_2(A) \leq \eta_1(A^\varepsilon) + \varepsilon \quad \forall A \subseteq \mathbb{T} \text{ medible}\},$$

en que A^ε corresponde a los $y \in \mathbb{T}$ tales que $\exists x \in A$ con $d_{\mathbb{T}}(x, y) < \varepsilon$. Es directo verificar que

$$\eta_1(A) \leq \eta_2\left(A^{\frac{2}{n}}\right) \;\wedge\; \eta_2(A) \leq \eta_1\left(A^{\frac{2}{n}}\right),$$

luego $d_(\eta_1, \eta_2) \leq 2k/n \leq 2p/n$ y tomando n grande se concluye el resultado.

□

La proposición anterior muestra que para n grande, los elementos dentro de un mismo conjunto $H \in \mathcal{H}_n$ tienen tasas muy parecidas, y por lo tanto podemos elegir a un elemento *representententante* de cada H.

Definición 5.6 *Para cada $n \in \mathbb{N}$ se define la secuencia de rasgos representantes de $\mathcal{D}_n$,*

$$\{w_D\}_{D\in\mathcal{D}_n} \subseteq \mathbb{T},$$

en que $w_D \in D$ para todo $D \in \mathcal{D}_n$. A partir de estos rasgos definimos la secuencia

$$\{\zeta_H\}_{H\in\mathcal{H}_n} \subseteq \mathcal{A}^{-0},$$

de configuraciones representantes, en que si H está definido por los conjuntos $D_1, D_2, \ldots, D_k$ entonces

$$\zeta_H \ := \ \delta_{w_{D_1}} + \delta_{w_{D_2}} + \cdots + \delta_{w_{D_k}}.$$

A continuación construiremos las tasas de transición de los procesos $\{\hat{W}^n\}_{n\in\mathbb{N}}$ utilizando la definición anterior.

Definición 5.7 *Para cada $n \in \mathbb{N}$ definimos las funciones b_n, m_n y λ_n a partir de b, m, λ y las secuencias $\{w_D\}_{D\in\mathcal{D}_n}$, $\{\zeta_H\}_{H\in\mathcal{H}_n}$ como*

$$b_n(y,\eta) := b(w_D, \zeta_H), \ \ m_n(y,\eta) := m(w_D,\zeta_H) \ \ y \ \ \lambda_n(y,\eta) := \lambda(w_D,\zeta_H),$$

en que D y H son tales que $y \in D$ y $\eta \in H$. De igual forma, tomando $n \in \mathbb{N}$ fijo y dados $y, z \in \mathbb{T}$ se define $g_n(y,z)$ como

$$g_n(y,z) \ := \ g(w_D, z),$$

donde D es tal que $y \in D$.

Es directo ver que las funciones b_n, m_n, λ_n y g_n de la definición anterior son estrictamente positivas y localmente acotadas, en que las cotas son las mismas que las de b, m y λ, y cumplen las hipótesis (H) si b, m, λ y g las cumplen. Utilizando tales funciones podemos construir finalmente los procesos $\{\hat{W}^n\}_{n\in\mathbb{N}}$ como sigue.

Definición 5.8 *Para cada $n \in \mathbb{N}$ se define el proceso W^n como uno construido a partir de la variable aleatoria Y_0 y las funciones b_n, m_n, λ_n y g_n como en las Definiciones 2.9 y 2.11. A partir de W se define $\hat{W}^n$ análogamente a $\hat{Y}$ como el proceso*

$$\hat{W}^n_t \ := \ W^n_{t\wedge\mathcal{T}_{p+1}}$$

a valores en $\mathcal{A}_{\leq p} \cup \{0\}$.

Las funciones introducidas en la Definición 5.7 no son necesariamente continuas, sin embargo ya habíamos mencionado que para los resultados que no involucran existencia de q.s.d. basta con asumir que las tasas son positivas y acotadas, luego $\hat{W}^n$ presenta las mismas propiedades que $\hat{Y}$ (no posee explosiones, es μ-s.s.i., etc.) salvo que no es un proceso de Feller.

Utilizaremos la notación $\hat{G}_n$, $\hat{Q}_n$ y $\hat{P}^n_t$ para referirnos al generador infinitesimal, kernel de transición y semigrupo de transición de $\hat{W}^n$ respectivamente, en que para los dos primeros se obtienen las expresiones;

$$
\begin{aligned}
\hat{G}_n f(\eta) = & \sum_{j=1}^{k} b(w_{D_j}, \zeta_H)[f(\eta+\delta_{y_j}) - f(\eta)] + \sum_{j=1}^{k} \lambda(w_{D_j}, \zeta_H)[f(\eta-\delta_{y_j}) - f(\eta)] \\
& + \sum_{j=1}^{k} m(w_{D_j}, \eta_H) \int_{\mathbb{T}} [f(\eta+\delta_z) - f(\eta)] g(w_{D_j}, z) d\sigma(z),
\end{aligned} \tag{5.1}
$$

$$
\hat{Q}_n(\eta, B) := \sum_{\substack{1\le j\le k \\ \eta+\delta_{y_j}\in B}} b(w_{D_j}, \zeta_H) \;+\; \sum_{\substack{1\le j\le k \\ \eta-\delta_{y_j}\in B}} \lambda(w_{D_j}, \zeta_H) \;+\; \sum_{j=1}^{k} m(w_{D_j}, \zeta_H) \int_{\eta+\delta_z\in B} g(w_{D_j}, z) d\sigma(z),
$$

en que $\eta \in H$ y $D_1, D_2, \ldots, D_k \in \mathcal{D}_n$ son tales que para $\vec{\eta} = \{y_1, y_2, \ldots, y_k\}$ se tiene que cada $y_j \in D_j$. Utilizaremos esta expresión del kernel de transición para mostrar que de cierta forma $\hat{W}^n$ toma valores sobre $\mathcal{H}_n$ que es un conjunto numerable.

Proposición 5.9 *Sea $n \in \mathbb{N}$ cualquiera, entonces para todo par de conjuntos $H, H' \in \mathcal{H}_n$ con $H, H' \subseteq \mathcal{A}_{\le p}$, todo par de configuraciones $\eta_1, \eta_2 \in H$ y todo $t \ge 0$ se cumple*

$$
\mathbb{P}_{\eta_1}(\hat{W}_t^n \in H') = \mathbb{P}_{\eta_2}(\hat{W}_t^n \in H').
$$

Demostración. El resultado se deduce fácilmente al comprobar que para η_1, η_2 perteneciente al mismo H se cumple

$$
\hat{Q}_n(\eta_1, H') = \hat{Q}_n(\eta_2, H'), \tag{5.2}
$$

con $H' \in \mathcal{H}_n$. En efecto, asumiendo cierta (5.2) mostraremos por inducción que para todo $m \ge 0$ y para todo H, H', η_1, η_2, t como en el enunciado se tiene

$$
\mathbb{P}_{\eta_1}\left(\hat{W}_t^n \in H',\ N_t = m\right) = \mathbb{P}_{\eta_2}\left(\hat{W}_t^n \in H',\ N_t = m\right).
$$

Para el caso $m = 0$, si $H \neq H'$ la probabilidad es nula comenzando tanto en η_1 como en η_2, y si $H = H'$ entonces

$$
\mathbb{P}_{\eta_1}\left(\hat{W}_t^n \in H',\ N_t = 0\right) \;=\; e^{-\hat{Q}_n(\eta_1)t} \;=\; \mathbb{P}_{\eta_1}\left(\hat{W}_t^n \in H',\ N_t = 0\right),
$$

donde hemos utilizado que $\hat{Q}_n(\eta_1) = \hat{Q}_n(\eta_2)$ gracias a (5.2). Supongamos ahora que el resultado se tiene para m, luego para probar que también se cumple en el caso $m+1$ tomemos $t \ge 0$, $H, H' \in \mathcal{H}_n$ y $\eta_1, \eta_2 \in H$, y utilizando la propiedad de Markov se deduce

$$
\begin{aligned}
\mathbb{P}_{\eta_1}\left(\hat{W}_t^n \in H',\ N_t = m+1\right) &= \int_0^t e^{-\hat{Q}(\eta_1)s} \sum_{H''\in\mathcal{H}_n} \mathbb{P}_{\zeta_{H''}}\left(\hat{W}_{t-s}^n \in H',\ N_{t-s} = m\right) \hat{Q}_n(\eta_1, H'') ds \\
&= \int_0^t e^{-\hat{Q}(\eta_2)s} \sum_{H''\in\mathcal{H}_n} \mathbb{P}_{\zeta_{H''}}\left(\hat{W}_{t-s}^n \in H',\ N_{t-s} = m\right) \hat{Q}_n(\eta_2, H'') ds \\
&= \mathbb{P}_{\eta_2}\left(\hat{W}_t^n \in H',\ N_t = m+1\right),
\end{aligned}
$$

pues por hipótesis de inducción la función $\eta \to \mathbb{P}_{\eta}\left(\hat{W}_{t-s}^n \in H',\ N_{t-s} = m\right)$ es constante sobre cada conjunto $H'' \in \mathcal{H}_n$.

Para mostrar que se cumple (5.2) tomemos $\eta_1, \eta_2 \in H$ con $\vec{\eta_1} = (x_1, x_2, \ldots, x_k)$ y $\vec{\eta_2} = (y_1, y_2, \ldots, y_k)$, luego para cada $j \leq k$ se tiene $x_j, y_j \in D_j$ con $D_j \in \mathcal{D}_n$, y entonces

$$b_n(x_j, \eta_1) \;=\; b(w_{D_j}, \zeta_H) \;=\; b_n(y_j, \eta_2),$$

y lo mismo ocurre para las tasas m y λ. Como $g(x_j, z) = g(w_{D_j}, z) = g(y_j, z)$, para deducir (5.2) basta notar que

- cada par x_j, y_j pertenece al mismo D_j, luego, como $\eta_1, \eta_2 \in H$ el par $\eta_1 + \delta_{x_j}, \eta_2 + \delta_{y_j}$ pertenecen a la misma clase de equivalencia en $\mathcal{H}_n$, y en particular

$$\eta_1 + \delta_{z_{l,1}} \in H' \iff \eta_2 + \delta_{z_{l,2}} \in H';$$

- de igual forma, como $x_j, y_j \in D_j$ y $\eta_1, \eta_2 \in H$ se tiene

$$\eta_1 - \delta_{z_{l,1}} \in H' \iff \eta_2 - \delta_{z_{l,2}} \in H';$$

- tomando $z \in \mathbb{T}$, como $\eta_1, \eta_2 \in H$ las configuraciones $\eta_1 + \delta_z$ y $\eta_2 \in \delta_z$ pertenecen a la misma clase de equivalencia en $\mathcal{H}_n$, luego

$$\eta_1 + \delta_z \in H' \iff \eta_2 + \delta_z \in H'.$$

Finalmente podemos concluir

$$\begin{aligned}
\hat{Q}_n(\eta_1, H') &= \sum_{\substack{1\leq j\leq k \\ \eta_1+\delta_{x_j}\in H'}} b(w_{D_j}, \zeta_H) + \sum_{\substack{1\leq j\leq k \\ \eta_1-\delta_{x_j}\in H'}} \lambda(w_{D_j}, \zeta_H) + \sum_{j=1}^{k} m(w_{D_j}, \zeta_H) \int_{\eta_1+\delta_z\in H'} g(w_{D_j}, z)d\sigma(z) \\
&= \sum_{\substack{1\leq j\leq k \\ \eta_2+\delta_{y_j}\in H'}} b(w_{D_j}, \zeta_H) + \sum_{\substack{1\leq j\leq k \\ \eta_2-\delta_{y_j}\in H'}} \lambda(w_{D_j}, \zeta_H) + \sum_{j=1}^{k} m(w_{D_j}, \zeta_H) \int_{\eta_2+\delta_z\in H'} g(w_{D_j}, z)d\sigma(z) \\
&= \hat{Q}_n(\eta_2, H').
\end{aligned}$$

□

5.3. Vectores propios maximales de $\hat{W}^n$

Ya se había mencionado que para $\hat{W}^n$ las tasas de transición no son continuas y por lo tanto no posee la propiedad de Feller. Debido a que el semigrupo no conserva la estructura topológica no es posible utilizar la compacidad del espacio para garantizar la existencia de distribuciones cuasi-estacionarias como en la Proposición 3.4. Sin embargo, la Proposición 5.9 permitirá adaptar el principal resultado obtenido en [21], y que postula la existencia, unicidad y convergencia exponencial a una distribución cuasi-estacionaria para procesos a espacio de estados discretos que vuelven rápidamente desde infinito.

Teorema 5.1 *[21, Teorema 1] Sea $n \in \mathbb{N}$ fijo, para $\hat{W}^n$ existe una única q.s.d. π_n y además existe $0 < \gamma < 1$ tal que para toda medida de probabilidad ν sobre $\mathcal{A}_{\leq p}$ y $t \geq 1$ se cumple*

$$\|\mathbb{P}_\nu(\hat{W}^n_t \in \cdot \mid \mathcal{T} > t) - \pi\|_{TV} \;\leq\; 2\gamma^{t-1},$$

en que $\|\cdot\|_{TV}$ es la norma de variación total y $\mathcal{T}$ es el tiempo de absorción de $\hat{W}^n$.

Demostración. Al igual que lo hecho en [21] mostraremos primero que existe una medida no negativa y no idénticamente nula ρ tal que para todo $\eta \in \mathcal{A}_{\leq p}$ y $s \geq 1$ se cumple

$$\mathbb{P}_\eta(\hat{W}_1^n \in \cdot \mid \mathcal{T} > s) \ \geq \ \rho. \tag{5.3}$$

Para ello probaremos primero las siguientes propiedades:

- existe $0 < \alpha < 1$ tal que $\forall \eta \in \mathcal{A}^{-0}$, $k \leq p$ y $A \subseteq \mathcal{A}_k$ medible se cumple
$$\mathbb{P}_\eta(\hat{W}_1^n \in A) \ \geq \ \alpha\mu(A); \tag{5.4}$$
- existe una constante $\kappa > 0$ tal que para todo $\eta \in \mathcal{A}_{\leq p}$ y todo $t \geq 0$ se cumple
$$\frac{\mathbb{P}_\eta(\mathcal{T} > t)}{\sup_{\eta'' \in \mathcal{A}_{\leq p}} \mathbb{P}_{\eta''}(\mathcal{T} > t)} \ \geq \ \kappa. \tag{5.5}$$

Para mostrar (5.4) supondremos por comodidad que los elementos en A no poseen rasgos repetidos. Tomemos $\vec{\eta} = (y_1, y_2, \ldots, y_l)$, luego, si consideramos los tiempos de salto $\tau_1, \tau_2, \ldots, \tau_{k+l}$ del proceso, la probabilidad del evento $\hat{W}^n_{\tau_{l+k}} \in A$ es mayor que la probabilidad del evento en que;

1. los primeros $l-1$ saltos corresponden a la muerte de individuos de η;
2. ocurre una mutación agregando un rasgo presente en $\overline{A}$;
3. muere el último individuo original de η;
4. luego de $k-1$ mutaciones sucesivas se llega hasta el conjunto A.

La probabilidad de este evento es sencilla de calcular, obteniéndose

$$\prod_{j=1}^{l-1}\left(\frac{\lambda_n(y_j, \eta_{j-1})}{\hat{Q}_n(\eta_{j-1})}\right) \int_{\overline{A}} \left(\prod_{j=2}^{k} \Theta_j\right) \frac{\lambda_n(y_l, \eta_1' + \delta_{y_l})}{\hat{Q}_n(\eta_1' + \delta_{y_l})} \frac{m_n(y_l, \eta_l)}{\hat{Q}_n(\eta_l)} g_n(y_l, z_1) \prod_{j=1}^{k} d\sigma(z_j), \tag{5.6}$$

donde hemos utilizado la abreviación

$$\Theta_j = \left(\frac{m_n(z_j, \eta_{j-1}')}{\hat{Q}_n(\eta_{j-1}')} g_n(z_{j-1}, z_j)\right),$$

y en que $\eta_0 = \eta$, $\eta_j = \eta_{j-1} - \delta_{y_j}$, y $\eta_j' = \sum_{r=1}^{j} \delta_{z_r}$. Tomando ahora $\overline{Q}$ una cota superior de Q en $\mathcal{A}_{\leq p}$ y $\underline{\lambda}, \underline{m}$ y $\underline{g}$ cotas inferiores positivas para λ, m y g, podemos acotar (5.6) inferiormente por

$$\left(\underline{\lambda}/\overline{Q}\right)^l \left(\underline{mg}/\overline{Q}\right)^k \sigma^k(\overline{A}).$$

De la definición de μ y recordando que $l \leq p$, existe $\alpha' > 0$ independiente de η con

$$\mathbb{P}_\eta(\hat{W}^n_{\tau_{l+k}} \in A) \ \geq \ \alpha'\mu(A),$$

luego se tiene

$$\mathbb{P}_\eta(\hat{W}_t^n \in A) \ \geq \ \mathbb{P}_\eta(\hat{W}^n_{\tau_{l+k}} \in A \ \wedge \ N_t = l+k) \ \geq \ \alpha'\mu(A)\mathbb{P}_\eta(N_t = l+k),$$

y tomando V_1, V_2 variables aleatorias con $V_1 \sim \Gamma(2p, \overline{Q})$ y $V_2 \sim exp(\underline{Q})$, es sencillo verificar

$$\mathbb{P}_\eta(N_t = l+k) \ \geq \ \mathbb{P}(V_1 \leq 1)\mathbb{P}(V_2 > 1) > 0,$$

en que la expresión anterior no depende de η, luego se concluye la propiedad.

Para mostrar (5.5) sea $H' \subseteq \mathcal{A}_{\leq p}$ un elemento de $\mathcal{H}_n$ y tomemos $\eta \in H'$. Para $t \geq 0$ se tiene

$$\mathbb{P}_{\eta'}(\mathcal{T} > t) \ \geq \ \mathbb{P}_{\eta'}(\mathcal{T} > t+1) \ = \ \int_{\mathcal{A}_{\leq p}} \mathbb{P}_{\eta}(\mathcal{T} > t)\mathbb{P}_{\eta'}(\hat{W}_1^n \in d\eta),$$

y por otro lado,

$$\mathbb{P}_{\eta}(\mathcal{T} > t) \ = \ \mathbb{P}_{\eta}(\hat{W}_t^n \in \mathcal{A}_{\leq p}) \ = \ \sum_{\substack{H \subseteq \mathcal{A}_{\leq p} \\ H \in \mathcal{H}_n}} \mathbb{P}_{\eta}(\hat{W}_t^n \in H),$$

luego gracias a la Proposición 5.9, $\mathbb{P}_{\eta}(\mathcal{T} > t)$ es una función constante sobre los elementos de $\mathcal{H}_n$. Podemos acotar entonces tal función de la forma

$$\begin{aligned} \mathbb{P}_{\eta'}(\mathcal{T} > t) \ &\geq \ \int_{\mathcal{A}_{\leq p}} \mathbb{P}_{\eta}(\mathcal{T} > t)\mathbb{P}_{\eta'}(\hat{W}_1^n \in d\eta) \\ &= \ \sum_{\substack{H \subseteq \mathcal{A}_{\leq p} \\ H \in \mathcal{H}_n}} \mathbb{P}_{\zeta_H}(\mathcal{T} > t)\mathbb{P}_{\eta'}(\hat{W}_1^n \in H), \end{aligned}$$

en que la suma es finita, luego existe H_0 tal que para todo $H \in \mathcal{H}_n$ con $H \subseteq \mathcal{A}_{\leq p}$,

$$\mathbb{P}_{\eta'}(\hat{W}_1^n \in H) \ \geq \ \mathbb{P}_{\eta'}(\hat{W}_1^n \in H_0).$$

Notemos que de la Proposición 5.9, nuevamente la elección de H_0 no depende de la configuración η', luego para todo $\eta'' \in \mathcal{A}_{\leq p}$ deducimos

$$\mathbb{P}_{\eta'}(\mathcal{T} > t) \ \geq \ \mathbb{P}_{\eta'}(\hat{W}_1^n \in H_0) \sum_{\substack{H \subseteq \mathcal{A}_{\leq p} \\ H \in \mathcal{H}_n}} \mathbb{P}_{\zeta_H}(\mathcal{T} > t) \ \geq \ \mathbb{P}_{\eta'}(\hat{W}_1^n \in H_0)\mathbb{P}_{\eta''}(\mathcal{T} > t), \tag{5.7}$$

en que la última desigualdad se tiene pues existe $H'' \in \mathcal{H}_n$ tal que $\eta'' \in H''$, y entonces $\mathbb{P}_{\eta''}(\mathcal{T} > t) = \mathbb{P}_{\zeta_{H''}}(\mathcal{T} > t)$. Definimos finalmente

$$\kappa \ := \ \min_{\substack{H \subseteq \mathcal{A}_{\leq p} \\ H \in \mathcal{H}_n}} \ \min_{\eta' \in \mathcal{A}_{\leq p}} \mathbb{P}_{\eta'}(\hat{W}_1^n \in H),$$

que es positivo gracias a que hay finitos conjuntos H y finitos valores que puede tomar tal función para cada H. Acotando por este valor la ecuación (5.7) se deduce el resultado.

Para encontrar la medida ρ de la desigualdad (5.3) tomemos $\eta \in \mathcal{A}_{\leq p}$, $A \subseteq \mathcal{A}_{\leq p}$ medible y $s \geq 1$, luego utilizando la propiedad de Markov y las desigualdades (5.4) y (5.5),

$$\begin{aligned} \mathbb{P}_{\eta}(s < \mathcal{T})\mathbb{P}_{\eta}(\hat{W}_1^n \in A \mid s < \mathcal{T}) \ &= \ \mathbb{P}_{\eta}(\hat{W}_1^n \in A, \ s < \mathcal{T}) \\ &= \ \int_A \mathbb{P}_{\eta'}(s-1 < \mathcal{T})\mathbb{P}_{\eta}(\hat{W}_1^n \in d\eta') \\ &\geq \ \kappa \, \mathbb{P}_{\eta}(\hat{W}_1^n \in A) \sup_{\eta'' \in \mathcal{A}_{\leq p}} \mathbb{P}_{\eta''}(s-1 < \mathcal{T}) \\ &\geq \ \kappa\alpha\mu(A)\mathbb{P}_{\eta}(s < \mathcal{T}), \end{aligned}$$

luego basta tomar $\rho \equiv \alpha\kappa\mu$, que cumple con (5.3). Ya definida la medida ρ, el resto de la demostración es idéntico a lo hecho en [21], por lo que presentamos aquí lo hecho en ese trabajo sin entrar en demasiados detalles;

1. se define la familia de kernels $\{R^T_{t,s}\}_{t\leq s\leq T}$ como $R^T_{t,s} := \mathbb{P}_\eta(\hat{W}^n_{t-s} \in \cdot \mid T-t < \mathcal{T})$. Utilizando sucesivamente la propiedad de Markov se puede demostrar que para todo $t \leq s \leq u \leq T$ se cumple $R^T_{t,s}R^T_{s,u} = R^T_{t,u}$;

2. para todo $s \geq 1$ y toda medida ν de probabilidad se tiene $\mathbb{P}_\nu(\hat{W}^n_1 \in \cdot \mid s < \mathcal{T}) \geq \rho$. En particular, la medida
$$\nu R^{t+s}_{t,t+1} - \rho \;:=\; \mathbb{P}_\nu(\hat{W}^n_1 \in \cdot \mid s < \mathcal{T}_0) - \rho,$$
es no negativa y al evaluarla sobre $\mathcal{A}_{\leq p}$ obtenemos $\gamma := 1-\rho(\mathcal{A}_{\leq p})$ donde este valor es estrictamente menor a 1, pues ρ no es idénticamente nula. Tomando ν_1, ν_2 dos medidas de probabilidad ortogonales y $s \leq T-1$ se tiene
$$\begin{aligned}\|\nu_1 R^T_{s,s+1} - \nu_2 R^T_{s,s+1}\|_{TV} &\leq \|\nu_1 R^T_{s,s+1} - \rho\|_{TV} + \|\nu_2 R^T_{s,s+1} - \rho\|_{TV}\\ &= 2\gamma \;\leq\; \gamma\|\nu_1-\nu_2\|_{TV};\end{aligned}$$

3. dadas dos medidas de probabilidad distintas ν_1 y ν_2 podemos separar su resta en la parte negativa y la parte positiva, que son ortogonales, luego
$$\|\nu_1 R^T_{s,s+1} - \nu_2 R^T_{s,s+1}\|_{TV} \leq\; \gamma\|\nu_1-\nu_2\|_{TV}.$$
Utilizando la desigualdad anterior y la propiedad de semigrupo de $R^T_{s,t}$ podemos deducir inductivamente que
$$\begin{aligned}\|\nu_1 R^T_{0,T} - \nu_2 R^T_{0,T}\|_{TV} &= \|\nu_1 R^T_{0,T-1}R^T_{T-1,T} - \nu_2 R^T_{0,T-1}R^T_{T-1,T}\|_{TV}\\ &\leq \gamma\|\nu_1 R^T_{0,T-1} - \nu_2 R^T_{0,T-1}\|_{TV} \;\leq\; 2\gamma^{\lfloor T\rfloor},\end{aligned}$$
y como $\gamma < 1$, este último término es menor a $2\gamma^{T-1}$. De la definición de $R^T_{0,T}$ hemos probado entonces que para ν_1, ν_2 medidas de probabilidad cualesquiera se cumple
$$\|\mathbb{P}_{\nu_1}(\hat{W}^n_T \in \cdot \mid T < \mathcal{T}) - \mathbb{P}_{\nu_2}(\hat{W}^n_T \in \cdot \mid T < \mathcal{T})\|_{TV} \;\leq\; 2\gamma^{T-1}; \tag{5.7}$$

4. Finalmente, para $t \geq 1$, $s \geq 0$ y una medida de probabilidad ν cualquiera se tiene
$$\mathbb{P}_\nu(\hat{W}^n_{t+s} \in \cdot \mid t+s < \mathcal{T}) \;=\; \nu R^{t+s}_{0,t+s} \;=\; \nu R^{t+s}_{0,s}R^{t+s}_{s,t+s} \;=\; \mathbb{P}_{\nu R^{t+s}_{0,s}}(\hat{W}^n_t \in \cdot \mid t < \mathcal{T}),$$
y entonces, gracias a la desigualdad anterior,
$$\|\mathbb{P}_\nu(\hat{W}^n_t \in \cdot \mid t < \mathcal{T}) - \mathbb{P}_\nu(\hat{W}^n_{t+s} \in \cdot \mid t+s < \mathcal{T})\|_{TV} \;\leq\; 2\gamma^{t-1}.$$
Como $\gamma < 1$, la familia de medidas $\{\mathbb{P}_\nu(\hat{W}^n_t \in \cdot \mid t < \mathcal{T})\}_{t\geq 1}$ es de Cauchy, y como el espacio es completo, convergen a una medida π_n, la cual, de acuerdo a la definición 3.3 es una distribución cuasi-límite, y por lo tanto una q.s.d..

Reemplazando π_n en lugar de ν_2 en (5.7) y utilizando la definición de q.s.d.,
$$\|\mathbb{P}_\nu(\hat{W}^n_T \in \cdot \mid T < \mathcal{T}) - \pi_n\|_{TV} \;\leq\; 2\gamma^{T-1},$$
de donde además se deduce la unicidad de la distribución cuasi-estacionaria.

□

Es interesante notar que una vez encontrada la medida ρ se utilizó únicamente la propiedad de Markov del proceso $\hat{W}^n$ para deducir la existencia y unicidad de la distribución cuasi-estacionaria. Podríamos preguntarnos si es posible encontrar con el mismo método una medida ρ de este tipo para el proceso $\hat{Y}$. Para responder esto notemos que fue fundamental en la demostración poder encontrar una constante κ tal que

$$\mathbb{P}_{\eta_1}(t < \mathcal{T}) \geq \kappa \mathbb{P}_{\eta_2}(t < \mathcal{T}), \tag{5.8}$$

donde κ no depende de t, η_1 ni η_2. Este tipo de desigualdades puede mostrarse para procesos de salto con espacio de estados finito, sin embargo siendo $\mathcal{A}_{\leq p}$ un espacio no numerable no ha sido posible adaptar ninguna de tales demostraciones a este proceso. La existencia de un $\kappa > 0$ tal que se cumpla (5.8) sigue siendo una interrogante.

A partir del teorema anterior sabemos no sólo que existe una q.s.d. π_n para $\hat{W}^n$, sino también que esta distribución es única y que bajo cualquier medida de probabilidad inicial ν las medidas $\mathbb{P}_\nu(\hat{W}^n_t \in \cdot \mid t < \mathcal{T})$ convergen exponencialmente a π_n. A pesar de ello, nada sabemos aún sobre el parámetro θ_n asociado a π_n en el Teorema 3.3 o la existencia de alguna función f_n que sea θ_n-invariante. Con el fin de estudiar tales elementos introducimos la matriz M de la siguiente manera.

Definición 5.9 *Sea k' la cantidad de elementos de $\mathcal{H}_n$ que son subconjuntos de $\mathcal{A}_{\leq p}$ y $H_1, H_2, \ldots, H_{k'}$ una enumeración de tales elementos. Se define $M \in \mathbb{R}^{(k'+1)\times(k'+1)}$ como*

$$M_{i,j} := \begin{cases} \hat{Q}_n(\zeta_{H_i}, H_j) & si\ i,j \leq k'\ y\ i \neq j \\ -\hat{Q}_n(\zeta_{H_i}) & si\ i,j \leq k'\ y\ i = j \\ \hat{Q}_n(\zeta_{H_i}, 0) & si\ i \leq k'\ y\ j = k'+1 \\ 0 & si\ i = k'+1. \end{cases}$$

La matriz $-M$ sin su última fila y columna corresponde a una matriz de Minkowski no singular y por lo tanto posee un valor propio $\theta_n > 0$ con vector propio por la izquierda $\vec{q}$ y vector propio por la derecha $\vec{v}$, ambos no negativos (ver [13, Sección 2.5]). Además, como M cumple

$$\forall 1 \leq j \leq k'+1, \qquad \sum_{i=1}^{k'+1} M_{i,j} = 0,$$

esta matriz corresponde al generador infinitesimal de un proceso de Markov a valores en $\{1, 2, \ldots, k'+1\}$ y para el cual $\{k'+1\}$ es un estado cementerio. Como el espacio de estados es discreto el proceso posee la propiedad de Markov fuerte, luego gracias al lema 3.1 se tiene

$$\theta_n = \sum_{i \in CL_1(k+1)} \vec{q}_i M_{i,k'+1} \leq \sup_{1 \leq i \leq k} M_{i,k'+1},$$

y como $M_{i,k'+1} = Q(\eta_{n,H_i}, 0) \leq \overline{Q}$ se concluye

$$0 < \theta_n \leq \overline{Q}. \tag{5.9}$$

Utilizando el vector $\vec{v}$ construiremos a continuación una función θ_n-invariante para $\hat{W}^n$:

Definición 5.10 *Sea $\vec{v}$ el vector propio por la derecha de $-M$ asociado a θ_n con $H_1, H_2, \ldots, H_{k'}$ la enumeración de los elementos de $\mathcal{H}_n$ contenidos en $\mathcal{A}_{\leq p}$ utilizada en la Definición 5.9. Definimos la función $f_n : \mathcal{A}_{\leq p} \cup \{0\} \to \mathbb{R}$ como*

$$f_n(\eta) \ := \ \sum_{j=1}^{k'} v_j \mathbf{1}_{H_j}(\eta).$$

Proposición 5.10 *La función f_n construida en la definición anterior es θ_n-invariante para $\hat{W}^n$.*

Demostración. Como f_n puede tomar una cantidad finita de valores es una función acotada, y entonces aplicando la Proposición 5.3 a $\hat{W}^n$, el proceso

$$f(\hat{W}^n_t) - f(\hat{W}^n_0) - \int_0^t \hat{G}_n f(\hat{W}^n_s) ds,$$

es una martingala. Sea entonces $\eta \in \mathcal{A}_{\leq p}$, $\vec{\eta} = (y_1, y_2, \ldots, y_k)$ cualquiera y sea $1 \leq l \leq k'$ tal que $\eta \in H_l$, luego por definición de f_n se tiene

$$\begin{aligned} \hat{G}_n f_n(\eta) \ &= \ \sum_{r=1}^{k'} v_r \left(\sum_{j=1}^{k} b_n(y_j, \eta) \mathbf{1}_{H_r}(\eta + \delta_y) + \sum_{j=1}^{k} \lambda_n(y_j, \eta) \mathbf{1}_{H_r}(\eta - \delta_y) \right. \\ &\quad \left. + \sum_{j=1}^{k} m_n(y_j, \eta) \int_{\mathbb{T}} \mathbf{1}_{H_r}(\eta + \delta_z) g_n(y_j, z) d\sigma(z) \right) \ - \ \vec{v}_l \hat{Q}_n(\eta) \\ &= \ \sum_{r=1}^{k'} \vec{v}_r \hat{Q}_n(\eta, H_r) \ - \ \vec{v}_l \hat{Q}_n(\eta), \end{aligned}$$

y gracias a la Proposición 5.9 se tiene $\hat{Q}_n(\eta, H_r) = \hat{Q}_n(\zeta_{H_l}, H_r)$, luego, como $\hat{Q}_n(\eta, H_l) = 0$,

$$\begin{aligned} \hat{G}_n f_n(\eta) \ &= \ \sum_{r \neq l} \vec{v}_r \hat{Q}_n(\zeta_{H_l}, H_r) \ - \ \vec{v}_l \hat{Q}_n(\eta) \\ &= \ \sum_{r \neq l} \vec{v}_r \overline{M}_{l,r} \ + \ \vec{v}_l \overline{M}_{l,l} \ = \ (\overline{M} \cdot \vec{v})_l, \end{aligned}$$

pero $\vec{v}$ es vector propio de $-M$ y por lo tanto $\hat{G}_n f_n(\eta) = -\theta_n \vec{v}_l = -\theta_n f_n(\eta)$. Hemos probado entonces que

$$\hat{G}_n f_n \ \equiv \ -\theta_n f_n,$$

y utilizando la Proposición 3.2 deducimos que f_n es θ_n-invariante.

$\square$

Hemos encontrado así una función positiva, acotada y θ_n-invariante para $\hat{W}^n$. Notando que π_n obtenido en el Teorema 5.1 induce un vector propio por la izquierda para $-M$ es directo que su tasa de extinción es θ_n, y por lo tanto, aplicando la Proposición 4.10, el proceso $\hat{W}^n$ es θ_n-recurrente positivo con vectores propios maximales π_n y f_n. Es interesante que para π_n conocemos algunas propiedades, y para f_n conocemos un método de construcción explícito, y será de utilidad en la próxima sección tener una cota superior $\overline{Q}$ para θ_n, la que no depende de n.

5.4. Convergencia de los vectores propios maximales

En la presente sección mostraremos un criterio sencillo para determinar la convergencia de los vectores propios maximales de los procesos $\hat{W}^n$ a los de $\hat{Y}$. Aún más importante es que bajo tales condiciones existe una única distribución cuasi-estacionaria para $\hat{Y}$, la cual es absolutamente continua respecto a μ y el proceso es Λ_1-recurrente positivo. Para ello comenzamos probando un resultado sencillo que justifica llamar *aproximaciones de* $\hat{Y}$ a la familia $\{\hat{W}^n\}$.

Proposición 5.11 *Definamos para cada $h : \mathcal{A}_{\leq p} \to \mathbb{R}$ acotada, su norma $\|h\|_\infty$ como*

$$\|h\|_\infty \ := \ \sup_{\eta \in \mathcal{A}_{\leq p}} |h(\eta)|,$$

y para $n \in \mathbb{N}$ cualquiera notemos $\{\hat{P}_t^n\}_{t \geq 0}$ y $\{\hat{P}_t\}_{t \geq 0}$ los semigrupos de transición de los procesos $\hat{W}^n$ y $\hat{Y}$ respectivamente. Para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que para todo $n \geq N$ y $A \subseteq \mathcal{A}_{\leq p}$ medible se cumple

$$\|\hat{P}_t^n(\mathbf{1}_A) - \hat{P}_t(\mathbf{1}_A)\|_\infty \leq \varepsilon. \tag{5.10}$$

Demostración. Comencemos mostrando que tomando $t \geq 0$ fijo, para todo $\varepsilon > 0$ existe $K \in \mathbb{N}$ tal que para todo $n \in \mathbb{N}$ y $\eta \in \mathcal{A}_{\leq p}$,

$$\mathbb{P}_\eta \left(N_t^n \geq K \right) \ \leq \ \varepsilon, \tag{5.11}$$

en que hemos utilizado la notación $N_t^n := N_t^{\hat{W}^n}$. Para ello notemos que tal evento equivale a que el K-ésimo tiempo de salto τ_K sea menor a t, donde tal tiempo de salto corresponde a la suma de K variables aleatorias $X_j \sim exp(\hat{Q}_n(\hat{W}^n_{\tau_i}))$, y como $\hat{Q}_n$ está acotado inferiormente por $\underline{Q} > 0$, es directo ver que

$$\mathbb{P}_\eta \left(N_t^n \geq K \right) \ \leq \ e^{-\underline{Q}} \sum_{k=K}^{\infty} (\underline{Q}^k / k!),$$

que no depende de n ni de η. Como esta expresión tiende a 0 cuando $K \to \infty$, se concluye (5.11), y como las cotas para las tasas de transición son válidas para $\hat{Y}$ podemos extender el resultado a este proceso también.

Notemos ahora que para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que para todo $A \subseteq \mathcal{A}_{\leq p}$ medible y $n \geq N$ se cumple

$$\left\| \hat{Q}_n(\cdot, A) - \hat{Q}(\cdot, A) \right\|_\infty \leq \varepsilon. \tag{5.12}$$

En efecto, tomemos $\eta \in \mathcal{A}_{\leq p}$ con $\vec{\eta} = (y_1, y_2, \ldots, y_k)$ y $H \in \mathcal{H}_n$ tal que $\eta \in H$. Para este caso se tiene

$$\begin{aligned}
|\hat{Q}_n(\eta, A) - \hat{Q}(\eta, A)| \;\leq\; & \sum_{\substack{1 \leq j \leq k \\ \eta + \delta_{y_j} \in A}} |b(w_{D_j}, \zeta_H) - b(y_j, \eta)| + \sum_{\substack{1 \leq j \leq k \\ \eta - \delta_{y_j} \in A}} |\lambda(w_{D_j}, \zeta_H) - \lambda(y_j, \eta)| \\
& + \sum_{j=1}^{k} |m(w_{D_j}, \zeta_H) - m(y_j, \eta)| \int_{\eta + \delta_z \in A} g(w_{D_j}, z) d\sigma(z) \\
& + \sum_{j=1}^{k} m(y_j, \eta) \int_{\eta + \delta_z \in A} |g(w_{D_j}, z) - g(y_j, z)| d\sigma(z),
\end{aligned}$$

donde $y_j \in D_j$ para cada j. Gracias a la Proposición 5.8, n suficientemente grande podemos acotar la expresión anterior de la forma

$$\begin{aligned}
|\hat{Q}_n(\eta, A) - \hat{Q}(\eta, A)| \;&\leq\; 2k\varepsilon + \sum_{j=1}^{k} \left(\int_{\eta + \delta_z \in A} \varepsilon g(w_{D_j}, z) d\sigma(z) + m(y_j, \eta) \int_{\eta + \delta_z \in A} \varepsilon d\sigma(z) \right) \\
&\leq\; (3 + \overline{m}) k\varepsilon.
\end{aligned}$$

en que $\overline{m}$ es una cota para m en $\mathcal{A}_{\leq p}$, de donde se concluye (5.12). Para demostrar (5.10) tomemos $\varepsilon > 0$ y $\eta \in \mathcal{A}_{\leq p}$, luego gracias a (5.11) se tiene

$$\begin{aligned}
\left| \hat{P}_t^n(\mathbf{1}_A)(\eta) - \hat{P}_t(\mathbf{1}_A)(\eta) \right| \;&=\; \left| \mathbb{P}_\eta(\hat{W}_t^n \in A) - \mathbb{P}_\eta(\hat{Y}_t \in A) \right| \\
&\leq\; \sum_{j=0}^{K-1} \left| \mathbb{P}_\eta(\hat{W}_t^n \in A,\ N_t^n = j) - \mathbb{P}_\eta(\hat{Y}_t \in A,\ N_t = j) \right| + 2\varepsilon,
\end{aligned}$$

y definiendo $\mathcal{S}_j := \{s_1, s_2, \ldots, s_j,\ 0 \leq s_1 + \cdots + s_j \leq t\}$, para cada j se tiene

$$\mathbb{P}_\eta(\hat{W}_t^n \in A,\ N_t^n = j) = \int_{\mathcal{S}_j} \int_A \int_{\mathcal{A}_{\leq p}^{j-1}} e^{-\sum_{i=1}^{j} \hat{Q}_n(\eta_{i-1}) s_i} \hat{Q}_n(\eta, d\eta_1) \cdots \hat{Q}_n(\eta_{j-1}, d\eta_j) ds_1, \ldots ds_j,$$

y una expresión análoga se tiene para $\mathbb{P}_\eta(\hat{Y}_t \in A,\ N_t = j)$, luego de (5.12), para cada j podemos tomar n_j independientemente de A de forma tal que

$$\left| \mathbb{P}_\eta(\hat{W}_t^n \in A,\ N_t^n = j) - \mathbb{P}_\eta(\hat{Y}_t \in A,\ N_t = j) \right| \leq \varepsilon,$$

y como hay una cantidad finita de tales j, basta tomar el máximo de los n_j para concluir

$$\left| \hat{P}_t^n(\mathbf{1}_A)(\eta) - \hat{P}_t(\mathbf{1}_A)(\eta) \right| \leq (K-1)\varepsilon + 2\varepsilon.$$

□

Como consecuencia directa del resultado anterior y de que el semigrupo de transición es una contracción obtenemos la siguiente propiedad, que será de particular utilidad para el estudio de la convergencia de los vectores propios maximales de los procesos $\{\hat{W}^n\}_{n\in\mathbb{N}}$.

Corolario 5.1 *Para todo $\varepsilon > 0$ y $C > 0$ existe $N \in \mathbb{N}$ tal que para todo $n \geq N$ y toda función $h : \mathcal{A}_{\leq p} \to \mathbb{R}^+$ con $h \leq C$ se cumple*

$$\|\hat{P}_t^n(h) - \hat{P}_t(h)\|_\infty \leq \varepsilon.$$

Demostración. Gracias a la Proposición 5.11 el resultado se tiene para las funciones de la forma $h = \mathbf{1}_A$ con A medible. Sea entonces $h \leq C$ cualquiera y tomemos los conjuntos $A_{j,k}$ como $A_{j,k} := h^{-1}\left(\frac{C}{k}[j, j+1]\right)$, con los cuales definimos h' de la forma

$$h' := \frac{C}{k}\sum_{j=0}^{k-1} j\mathbf{1}_{A_{j,k}},$$

que para k grande cumple $\|h - h'\|_\infty \leq (\varepsilon/3)$, y entonces

$$\begin{aligned} \|\hat{P}_t^n(h) - \hat{P}_t(h)\|_\infty &\leq \|\hat{P}_t^n(h) - \hat{P}_t^n(h')\|_\infty + \|\hat{P}_t^n(h') - \hat{P}_t(h')\|_\infty + \|\hat{P}_t(h') - \hat{P}_t(h)\|_\infty \\ &\leq \|h - h'\|_\infty + C\sum_{j=0}^{k} \|\hat{P}_t^n(\mathbf{1}_{A_{j,k}}) - \hat{P}_t(\mathbf{1}_{A_{j,k}})\|_\infty + \|h - h'\|_\infty, \end{aligned}$$

donde hemos utilizado que $\|P_t^n(f)\|_\infty \leq \|f\|_\infty$, y tomando entonces n grande independientemente de los $A_{j,k}$, el término central es menor a $\varepsilon/3$ de donde se deduce que

$$\|\hat{P}_t^n(h) - \hat{P}_t(h)\|_\infty \leq \varepsilon.$$

□

A continuación se presentan los resultados de convergencia antes mencionados, de entre los cuales el primero es de carácter general pero no entrega mucha información respecto al comportamiento de $\hat{Y}$. El segundo resultado requiere una hipótesis adicional pero brinda no sólo la convergencia de los vectores propios maximales, sino además la unicidad de la q.s.d. para $\hat{Y}$ y la Λ^*-clasificación de este proceso.

Teorema 5.2 *Sea $\{\hat{W}^n\}_{n\in\mathbb{N}}$ la familia de procesos construidos como en la Definición 5.8, para los cuales, gracias al Teorema 5.1 existen únicas distribuciones cuasi-estacionarias $\{\pi_n\}_{n\in\mathbb{N}}$. Si consideramos $\{\theta_n\}_{n\in\mathbb{N}}$ los parámetros asociados a estas distribuciones y $\{f_n\}_{n\in\mathbb{N}}$ las funciones θ_n-invariantes construidas en la Definición 5.10, se cumple;*

1. *existe una distribución cuasi-estacionaria π para $\hat{Y}$ tal que $\pi_n \to \pi$ débilmente salvo subsucesión;*

2. *normalizando las funciones f_n de forma tal que*

$$\forall n \in \mathbb{N}, \quad \int_{\mathcal{A}_{\leq p}} f_n(\eta)u(\eta)d\mu(\eta) = 1 \tag{5.13}$$

para alguna función $u : \mathcal{A}_{\leq p} \to \mathbb{R}^+$ μ-integrable, si se cumple la condición

$$\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty \tag{5.14}$$

entonces se tienen los siguientes resultados:

- *existen $\theta \in \mathbb{R}$, una medida de probabilidad π y una función f tales que $\theta_n \to \theta$, $\pi_n \to \pi$ débilmente y $f_n \to f$ en la topología débil-* de $L^\infty(\mu)$;*
- *π es la única distribución cuasi-estacionaria del proceso $\hat{Y}$. Tal medida es equivalente a μ y tiene como parámetro asociado a θ;*
- *f es una función acotada y θ-invariante y el proceso es θ-recurrente positivo.*

Demostración. Para mostrar la primera parte del teorema notemos que como $\mathcal{A}_{\leq p}$ es compacto, la secuencia de medidas de probabilidad $\{\pi_n\}_{n\in\mathbb{N}}$ posee una subsucesión débilmente convergente a una medida de probabilidad π, esto es, para toda función h continua,

$$\langle \pi_{n_k}, h\rangle \xrightarrow[k\to\infty]{} \langle \pi, h\rangle, \tag{5.15}$$

y por otro lado, como cada θ_{n_k} cumple $0 < \theta_{n_k} \leq \overline{Q}$, esta sucesión converge a algún θ con $0 \leq \theta \leq \overline{Q}$ salvo subsucesión. Mostremos entonces que π es una q.s.d. para $\hat{Y}$. En virtud del Teorema 3.3 es suficiente con mostrar que para todo $t \geq 0$ y $h : \mathcal{A}_{\leq p} \to \mathbb{R}$ continua se cumple

$$\langle \pi, \hat{P}_t(h)\rangle \;=\; e^{-\theta t}\langle \pi, h\rangle, \tag{5.16}$$

y para ello notemos que para cada k, como π_{n_k} es q.s.d. para $\hat{W}^{n_k}$, se tiene

$$\begin{aligned}
\left|\langle \pi, \hat{P}_t(h)\rangle - e^{-\theta t}\langle \pi, h\rangle\right| \;&\leq\; \left|\langle \pi, \hat{P}_t(h)\rangle - \langle \pi_{n_k}, \hat{P}_t(h)\rangle\right| \;+\; \left|\langle \pi_{n_k}, \hat{P}_t(h)\rangle - \langle \pi_{n_k}, \hat{P}_t^{n_k}(h)\rangle\right| \\
&\quad+\; \left|e^{-\theta_{n_k} t}\langle \pi_{n_k}, h\rangle - e^{-\theta t}\langle \pi_{n_k}, h\rangle\right| \;+\; \left|e^{-\theta t}\langle \pi_{n_k}, h\rangle - e^{-\theta t}\langle \pi, h\rangle\right| \\
&\leq\; \left|\langle \pi, \hat{P}_t(h)\rangle - \langle \pi_{n_k}, \hat{P}_t(h)\rangle\right| \;+\; \left\|\hat{P}_t(h) - \hat{P}_t^{n_k}(h)\right\|_\infty \\
&\quad+\; \|h\|_\infty \left|e^{-\theta_{n_k} t} - e^{-\theta t}\right| \;+\; e^{-\theta t}\left|\langle \pi_{n_k}, h\rangle - \langle \pi, h\rangle\right|,
\end{aligned}$$

pero $\hat{P}_t$ posee la propiedad de Feller, luego $\hat{P}_t(h)$ es continua y entonces

$$\left|\langle \pi, \hat{P}_t(h)\rangle - \langle \pi_{n_k}, \hat{P}_t(h)\rangle\right| \xrightarrow[k\to\infty]{} 0.$$

Por otro lado, h es continua sobre un compacto, luego es acotada y entonces gracias al Corolario 5.1 se debe tener $\left\|\hat{P}_t(h) - \hat{P}_t^{n_k}(h)\right\|_\infty \xrightarrow[k\to\infty]{} 0$, luego se deduce que para $\varepsilon > 0$ tomando k suficientemente grande se tiene

$$\left|\langle \pi, \hat{P}_t(h)\rangle - e^{-\theta t}\langle \pi, h\rangle\right| \;\leq\; \varepsilon,$$

y como esto se cumple para todo $\varepsilon > 0$ se deduce el primer resultado.

Mostremos ahora que bajo (5.13) y (5.14) se tienen los resultados del enunciado. Para ello, sea π una distribución cuasi-estacionaria para $\hat{Y}$ (ya sabemos que $\hat{Y}$ posee al menos una de tales distribuciones), y tomemos θ como en la primera parte. Como $\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty$ sabemos en particular que la sucesión $\{f_n\}_{n\in\mathbb{N}}$ se encuentra acotada en $L^\infty(\pi)$, luego posee una subsucesión $\{f_{n_k}\}_{k\in\mathbb{N}}$ débil-* convergente a una cierta función $f \in L^\infty(\pi)$, esto es, para toda función $h \in L^1(\pi)$ se tiene

$$\int_{\mathcal{A}_{\leq p}} f_{n_k}(\eta)h(\eta)d\pi(\eta) \xrightarrow[k\to\infty]{} \int_{\mathcal{A}_{\leq p}} f(\eta)h(\eta)d\pi(\eta).$$

Es directo notar que f es π-c.s. no negativa, en efecto, como π es una medida de probabilidad las indicatrices son π-integrables y tomando $B_f := \{\eta \in \mathcal{A}_{\leq p},\ f(\eta) < 0\}$ se tiene

$$0 \ \leq\ \lim_{k\to\infty} \int_{\mathcal{A}_{\leq p}} f_{n_k}(\eta)\mathbf{1}_{B_f}(\eta)d\pi(\eta) \ =\ \int_{B_f} f(\eta)d\pi(\eta) \ \leq\ 0,$$

de donde necesariamente $\pi(B_f) = 0$. Por otro lado, de la Proposición 4.4 se desprende que $\mu << \pi$, luego, existe una función w positiva y π-integrable tal que $d\mu = wd\pi$. Utilizando la función w, (5.13) queda de la forma

$$\int_{\mathcal{A}_{\leq p}} f_n(\eta)u(\eta)w(\eta)d\pi(\eta) \ =\ 1,$$

y como u es μ-integrable, uw es π-integrable, de donde

$$1 \ =\ \lim_{k\to\infty} \int_{\mathcal{A}_{\leq p}} f_{n_k}(\eta)u(\eta)w(\eta)d\pi(\eta) \ =\ \int_{\mathcal{A}_{\leq p}} f(\eta)u(\eta)w(\eta)d\pi(\eta),$$

y entonces f no puede ser π-c.s. nula.

Mostraremos a continuación que f es θ-invariante para $\hat{Y}$, y para ello bastará mostrar que para toda función $h \in L^1(\pi)$ positiva y para todo $t \geq 0$ se cumple

$$\langle h, \hat{P}_t(f)\rangle_\pi \ =\ e^{-\theta t}\langle h, f\rangle_\pi, \tag{5.17}$$

en que para $h \in L^1(\pi)$ y $f \in L^\infty(\pi)$ hemos introducido la notación

$$\langle h, f\rangle_\pi \ :=\ \int_{\mathcal{A}_{\leq p}} h(\eta)f(\eta)d\pi(\eta).$$

Así, tomando $h \in L^1(\pi)$ fija y $k \in \mathbb{N}$ cualquiera podemos utilizar que f_{n_k} es θ_{n_k}-invariante para $\hat{W}^{n_k}$ y deducir

$$\begin{aligned}
\left|\langle h, \hat{P}_t(f)\rangle_\pi - e^{-\theta t}\langle h, f\rangle_\pi\right| \ &\leq\ \left|\langle h, \hat{P}_t(f)\rangle_\pi - \langle h, \hat{P}_t(f_{n_k})\rangle_\pi\right| + \left|\langle h, \hat{P}_t(f_{n_k})\rangle_\pi - \langle h, \hat{P}^{n_k}_t(f_{n_k})\rangle_\pi\right| \\
&\quad + \left|e^{-\theta_{n_k} t}\langle h, f_{n_k}\rangle_\pi - e^{-\theta t}\langle h, f_{n_k}\rangle_\pi\right| + \left|e^{-\theta t}\langle h, f_{n_k}\rangle_\pi - e^{-\theta t}\langle h, f\rangle_\pi\right| \\
&\leq\ \left|\langle h, \hat{P}_t(f)\rangle_\pi - \langle h, \hat{P}_t(f_{n_k})\rangle_\pi\right| + \|h\|_1 \left\|\hat{P}_t(f_{n_k}) - \hat{P}^{n_k}_t(f_{n_k})\right\|_\infty \\
&\quad + \|h\|_1 \sup_{n\in\mathbb{N}} \|f_n\|_\infty \left|e^{-\theta_{n_k} t} - e^{-\theta t}\right| + e^{-\theta t}\left|\langle h, f_{n_k}\rangle_\pi - \langle h, f\rangle_\pi\right|,
\end{aligned}$$

en que hemos utilizado la notación $\|\cdot\|_1$ y $\|\cdot\|_\infty$ para referirnos a las normas en $L^1(\pi)$ y $L^\infty(\pi)$ respectivamente. Gracias al Teorema 4.2 el semigrupo $\hat{P}_t$ es débil-* continuo como operador de $L^\infty(\pi)$ en sí mismo, luego tenemos

$$\left|\langle h, \hat{P}_t(f)\rangle_\pi - \langle h, \hat{P}_t(f_{n_k})\rangle_\pi\right| \xrightarrow[k\to\infty]{} 0,$$

y por otro lado, gracias al Corolario 5.1 el segundo término tiende a cero, y lo mismo con el resto de los términos debido a la selección de la subsucesión.

Hemos demostrado que existe una función $f \in L^\infty(\pi)$ no negativa y no idénticamente nula tal que $\pi - c.s.$ para todo $t \geq 0$ se cumple $\hat{P}_t(f) \;=\; e^{-\theta t} f$. Sea ahora $\theta' > 0$ la tasa de extinción de π y notemos que

$$\langle \pi, f\rangle \;=\; e^{\theta t}\langle \pi, \hat{P}_t(f)\rangle \;=\; e^{(\theta-\theta')t}\langle \pi, f\rangle,$$

pero como f no es π-c.s. nula y es acotada se debe tener $0 < \langle \pi', f\rangle < \infty$, y entonces $\theta = \theta'$. Utilizando la Proposición 4.11 deducimos que $\hat{Y}$ debe ser θ-recurrente positivo con vectores propios maximales π y f. En particular hemos mostrado que, de existir una distribución cuasi-estacionaria, ésta es un vector propio maximal del proceso y entonces, gracias al Teorema 4.7 existe una única q.s.d. la cual, en virtud del mismo teorema, es equivalente a μ.

Falta por ver entonces la convergencia de las secuencias $\{\theta_n\}_{n\in\mathbb{N}}$, $\{\pi_n\}_{n\in\mathbb{N}}$ y $\{f_n\}_{n\in\mathbb{N}}$. Para ello, notemos que en la demostración anterior se dedujo que si $\{\theta_n\}_{n\in\mathbb{N}}$ posee una subsucesión convergente, digamos a θ', entonces θ' es el parámetro asociado a los vectores propios maximales del proceso, y por lo tanto $\theta' = \Lambda^*$. Así, toda subsucesión convergente de $\{\theta_n\}_{n\in\mathbb{N}}$ converge a Λ^*, y como la sucesión es acotada debe tenerse

$$\theta_n \xrightarrow[n\to\infty]{} \Lambda^*.$$

De igual forma, en la primera parte de la demostración obtuvimos que si $\{\pi_n\}_{n\in\mathbb{N}}$ posee una subsucesión débilmente convergente, ésta converge a una distribución cuasi-estacionaria π para $\hat{Y}$, pero tal q.s.d. es única y por lo tanto, como $\{\pi_n\}_{n\in\mathbb{N}}$ vive en un conjunto débil-compacto se tiene

$$\pi_n \xrightarrow[n\to\infty]{} \pi,$$

en que la convergencia es en la topología débil. De manera análoga podemos deducir que

$$f_n \xrightarrow[n\to\infty]{} f$$

en que la convergencia es en la topología débil-* de $L^\infty(\mu)$.

□

Notemos que gracias al teorema anterior, bajo las condiciones adecuadas el proceso $\hat{Y}$ es Λ_1-recurrente positivo y su medida Λ_1-invariante π es acotada, luego, como es μ-s.s.i. podemos obtener del Teorema 4.9 que para todo $\eta \in \mathcal{A}_{\leq p}$ y $A \subseteq \mathcal{A}_{\leq p}$ medible,

$$\mathbb{P}_\eta(\hat{Y}_t \in A \mid t < \mathcal{T}) \xrightarrow[t\to\infty]{} \pi(A),$$

y entonces π es un límite de Yaglom.

5.5. Resultados numéricos

El Teorema 5.2 entrega un método que permite dar luces respecto a las propiedades de las distribuciones cuasi-estacionarias en el caso de un espacio de estados continuo. En efecto, a partir de la discretización del espacio de rasgos podemos construir las matrices M, luego, calculando los vectores propios asociados al valor propio minimal, y normalizándolos de alguna forma podemos calcular el valor máximo que toman. El teorema anterior muestra que estos valores máximos serán entonces indicadores de convergencia.

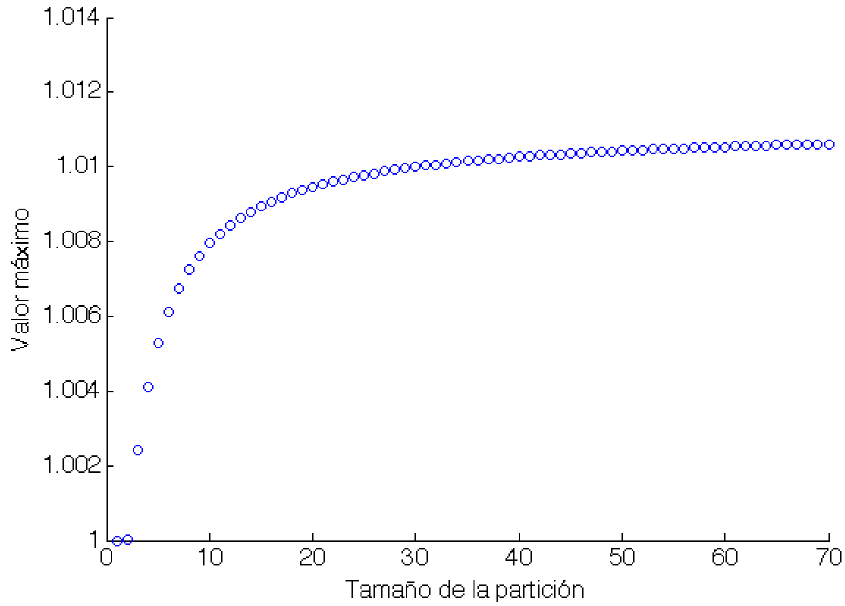
Con el fin de entender de forma algo más empírica lo que ocurre con estos indicadores se han estudiado algunos casos específicos, para los cuales hemos asumido:

1. el espacio de rasgos $\mathbb{T}$ será el conjunto $[0,1] \subseteq \mathbb{R}$ y la medida de probabilidad σ corresponde a la medida de Lebesgue sobre este intervalo;
2. para cada n hemos considerado $\mathcal{D}_n$ como la partición de $[0,1]$ en n intervalos de igual medida, esto es, todo $D \in \mathcal{D}_n$ es de la forma $[k/n, (k+1)/n)$ en que $0 \leq k \leq n-1$ y su rasgo representante w_D corresponde al punto medio del intervalo;
3. la normalización de los vectores propios fue elegida como $\int_{\mathcal{A}_{\leq p}} f_n(\eta) \mathbf{1}_{\mathcal{A}_1}(\eta) d\mu(\eta) = 1$, que en este caso, para cada n queda de la forma
$$\frac{1}{n} \sum_{D_k \subseteq \mathcal{A}_1} \vec{v}_k \ = \ 1;$$
4. debido al rápido crecimiento de las matrices M al aumentar el tamaño de la partición se decidió tomar $p = 3$, pues la construcción de tales matrices toma un tiempo de orden n^p en que n es el tamaño de la partición.

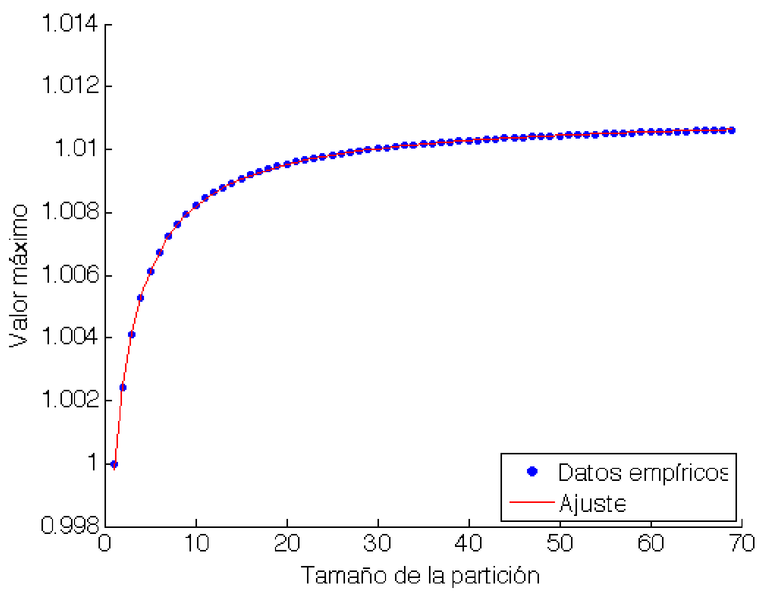
La diferencia entre los casos estudiados radica entonces en la elección de las tasas b, m y λ, junto con la elección de g. Para el primer ejemplo se utilizaron las siguientes funciones:

- $\lambda_w(\eta) \ = \ 4\|\eta\|;$
- $m_w(\eta) \ = \ \|\eta\|;$
- $b_w(\eta) \ = \ \begin{cases} w - w^2 + 1 & \text{si} \quad \|\eta\| = 1 \\ 4 & \text{si} \quad \|\eta\| = 2 \\ 6 & \text{si} \quad \|\eta\| = 3 \end{cases};$
- $g(a,b) \ = \ \dfrac{(a+1)e^{-(a+1)b}}{1 - e^{-(a+1)}}.$

Se calcularon los máximos de los vectores propios normalizados para un total de 70 matrices y se obtuvo como resultado el siguiente comportamiento de la sucesión.



En este caso, la secuencia m_n pareciera estabilizarse rápidamente cerca del valor 1.01, en efecto, realizando un ajuste en matlab para estos datos salvo el primero se obtiene la siguiente figura.

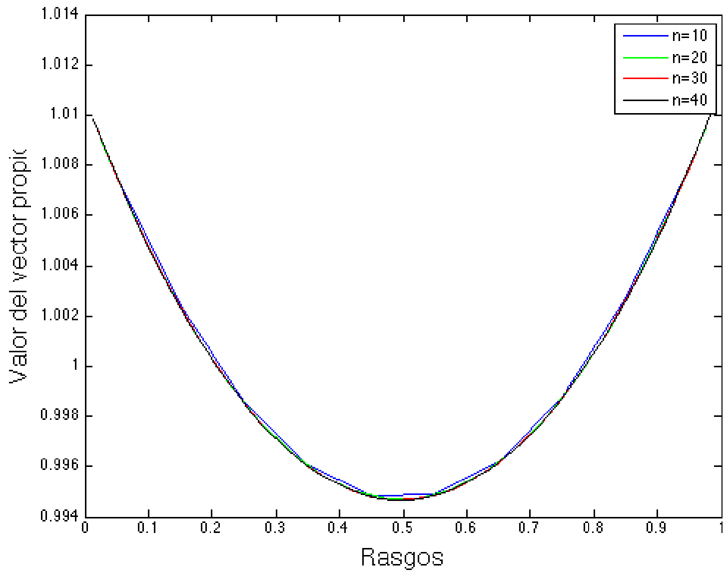


En que el ajuste corresponde a la función

$$\frac{1{,}011x + 2{,}161}{x + 2{,}173},$$

donde la suma de los errores al cuadrado es de $7{,}34 \times 10^{-8}$ y el valor de R^2 ajustado es de 0,9997. Podemos asumir entonces con un alto nivel de confianza que los valores máximos están acotados. En tal caso, del Teorema 5.2 podemos desprender entonces la unicidad de la distribución cuasi-estacionaria para el modelo y la Λ_1-positividad del proceso.

Graficando los vectores propios (interpolados) en $\mathcal{A}_1$ para distintos valores del tamaño de la partición de $\mathbb{T}$, n, se puede observar claramente su convergencia a una función suave.



La convergencia observada pareciera ser bastante rápida, y la función a la que tienden los vectores propios pareciera guardar relación directa con la forma de las tasas de transición.

Para ilustrar lo que ocurre al tomar tasas de transición de una mayor complejidad, se utilizaron para el segundo ejemplo las siguientes funciones, donde se asume siempre que los rasgos x, y, z cumplen $x \leq y \leq z$.

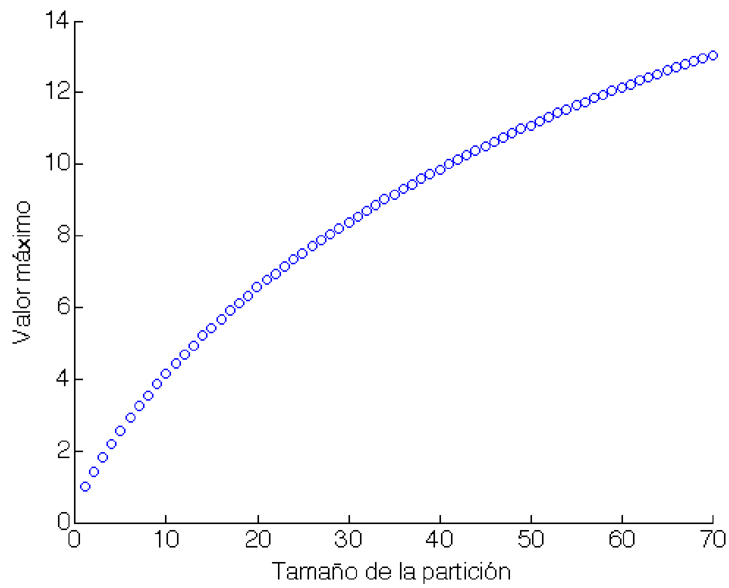
- $\lambda_w(\eta) = \begin{cases} 2^w & \text{si} \quad \eta = \delta_x \\ w + 2x + 3y + 2 & \text{si} \quad \eta = \delta_x + \delta_y \\ w + 2x + 3y - 4z + 5 & \text{si} \quad \eta = \delta_x + \delta_y + \delta_z \end{cases}$

- $m_w(\eta) = \begin{cases} w + 1 & \text{si} \quad \eta = \delta_x \\ w + x + y + 1 & \text{si} \quad \eta = \delta_x + \delta_y \\ 2^w + 3^x - y - 4^z + 5 & \text{si} \quad \eta = \delta_x + \delta_y + \delta_z \end{cases}$

- $b_w(\eta) = \begin{cases} 4w^2 + 4w + 1 & \text{si} \quad \eta = \delta_x \\ (w+1)2^{x-y} & \text{si} \quad \eta = \delta_x + \delta_y \\ w^2 + x^2 - y^2 - z^2 + 3 & \text{si} \quad \eta = \delta_x + \delta_y + \delta_z \end{cases}$

- $g(a, b) = \dfrac{1}{2}\left(\dfrac{z^{1+a}(1-z)^{2-a}}{\beta(2+a, 3-a)} + 1\right)$

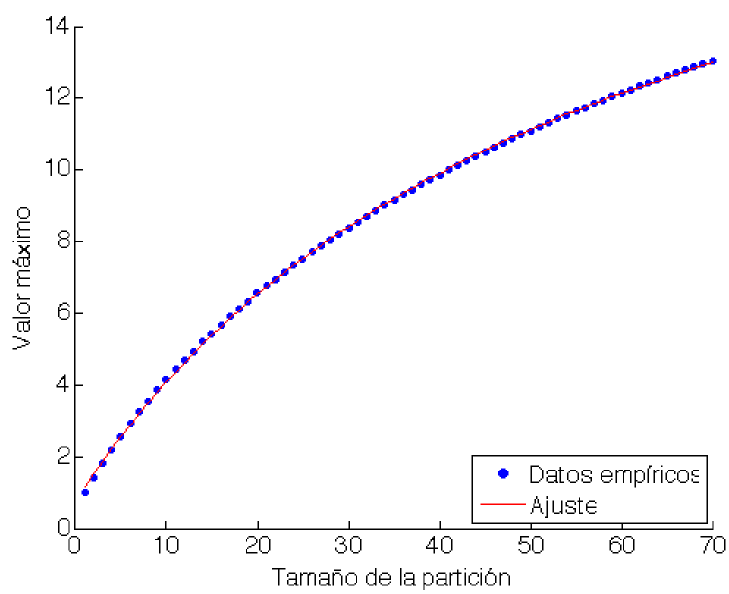
Al igual que en el ejemplo anterior se calcularon los máximos de los vectores propios normalizados para un total de 70 matrices, obteniéndose el siguiente comportamiento de la sucesión.



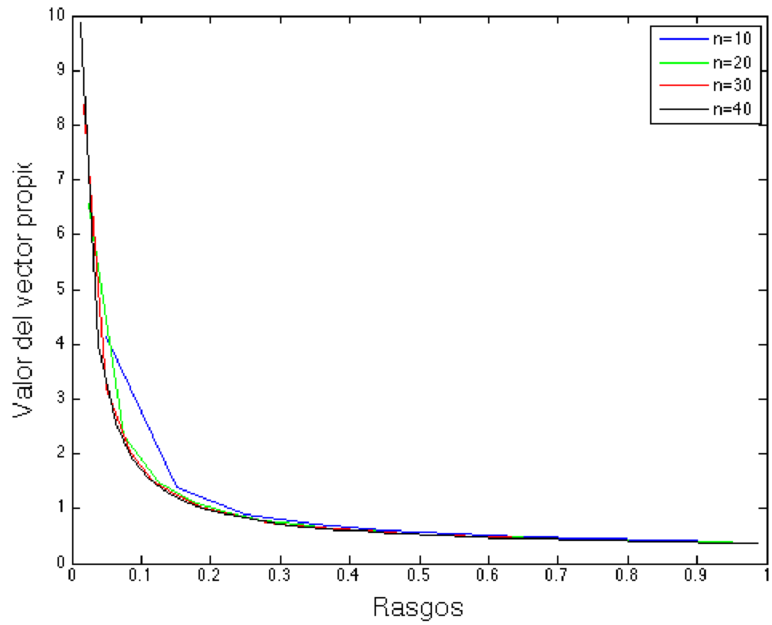
Se realizó una serie de ajustes para estos datos, de entre los cuales destacó la función

$$\frac{22{,}92x + 44{,}09}{x + 57},$$

obteniéndose en este caso una suma de errores cuadrados igual a 0,1423 y un valor de R^2 ajustado de 0,9998.



Aunque en este caso no se observa la convergencia de estos valores, graficando nuevamente los vectores propios sobre $\mathcal{A}_1$ para distintos tamaños de la partición se observa una tendencia clara.



Como puede observarse, los vectores nuevamente parecieran converger a una función f, la cual no sabemos si se mantiene acotada para valores cercanos a 0. Para estudiar este acotamiento se realizaron ajustes para los vectores propios en el caso de una partición de 30 elementos y en el de una partición de 40 elementos, obteniéndose valores de R^2 ajustado iguales a 1 para las funciones

$$\frac{0{,}2135x + 0{,}1562}{x + 0{,}0024} \quad \text{y} \quad \frac{0{,}2129x + 0{,}1553}{x + 0{,}0035},$$

que son ajustes para el caso de 30 y 40 elementos respectivamente. Notando que el valor que toma la segunda función en 0 es menor que el valor que toma la primera en tal punto, y que para valores pequeños una suerte de monotonía, no es ilógico pensar que existe convergencia.

5.6. El caso de tasas no continuas

Tal y como se mencionó en la sección 2.6, la continuidad de las tasas es una hipótesis que hemos utilizado casi en la totalidad de los casos para deducir cotas sobre $\mathcal{A}_{\leq p}$, salvo en los resultados de existencia de q.s.d. del capítulo 3. En este capítulo, sin embargo, no contar con la continuidad de las tasas significa perder la noción de *aproximación* de los procesos $\hat{W}^n$, pues la disminución del tamaño para los elementos en $\mathcal{D}_n$ no garantiza que podamos ver las tasas como constantes. Al trabajar con tasas no continuas será necesario introducir la Proposición 5.8 como un requisito para la construcción de $\mathcal{D}_n$ y $\mathcal{H}_n$.

Definición 5.11 *Diremos que la familia de particiones $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ inducida por $\{\mathcal{D}_n\}_{n\in\mathbb{N}}$ es una $\mathcal{A}_{\leq p}$-aproximación si para todo $\varepsilon > 0$ existe $n_\varepsilon \in \mathbb{N}$ tal que para $n \geq n_\varepsilon$,*

$$|b_{x_j}(\eta_1) - b_{y_j}(\eta_2)| + |m_{x_j}(\eta_1) - m_{y_j}(\eta_2)| + |\lambda_{x_j}(\eta_1) - \lambda_{y_j}(\eta_2)| + |g(x_j, z) - g(y_j, z)| \leq \varepsilon e^k$$

en que η_1, η_2 y las secuencias $\{x_i\}, \{y_i\}$ son como en la Proposición 5.8.

La elección de las particiones $\mathcal{D}_n$ que generan $\mathcal{A}_{\leq p}$-aproximaciones puede ser muy compleja, y a veces hasta imposible de obtener, sin embargo, para aquellos casos en los cuales existan, podemos construir los procesos $\hat{W}^n$ tal y como se hizo en la Definición 5.8 (para los cuales la totalidad de los resultados en la sección 5.3 se mantiene válida) y deducir los siguientes resultados.

Proposición 5.12 *Sea $h : \mathcal{A}_{\leq p} \to \mathbb{R}$ medible y acotada, entonces*

$$\|\hat{P}_t^n(h) - \hat{P}_t(h)\|_\infty \xrightarrow[n\to\infty]{} 0$$

Teorema 5.3 *Sea $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ una $\mathcal{A}_{\leq p}$-aproximación inducida por $\{\mathcal{D}_n\}_{n\in\mathbb{N}}$ y tomemos $\{\hat{W}^n\}_{n\in\mathbb{N}}$ la familia de procesos construidos como en la Definición 5.8, para los cuales existen únicas distribuciones cuasi-estacionarias $\{\pi_n\}_{n\in\mathbb{N}}$, las cuales son absolutamente continuas respecto a μ y por lo tanto poseen densidades $\{h_n\}_{n\in\mathbb{N}}$ respecto a tal medida. Si consideremos $\{\theta_n\}_{n\in\mathbb{N}}$ y $\{f_n\}_{n\in\mathbb{N}}$ como en el Teorema 5.2 entonces se cumple:*

1. *si existe algún $\beta > 1$ tal que*

   $$\limsup_{n\to\infty} \int_{\mathcal{A}_{\leq p}} [h_n(\eta)]^\beta \, d\mu_s(\eta) \ < \ \infty,$$

   en que μ_s es la medida definida en la Proposición 4.8, entonces existe una distribución cuasi-estacionaria π para $\hat{Y}$, la cual es absolutamente continua respecto a μ y tal que $\pi_n \to \pi$ débilmente salvo subsucesión.

2. *normalizando las funciones f_n de forma tal que*

   $$\forall n \in \mathbb{N}, \quad \int_{\mathcal{A}_{\leq p}} f_n(\eta)u(\eta)d\mu(\eta) \ = \ 1,$$

   para alguna función $u : \mathcal{A}_{\leq p} \to \mathbb{R}^+$ μ-integrable, si se cumple la condición

   $$\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty,$$

   entonces se tienen los siguientes resultados:
   - *existen $\theta \in \mathbb{R}$, una medida de probabilidad π y una función f tales que $\theta_n \to \theta$, $\pi_n \to \pi$ débilmente y $f_n \to f$ en la topología débil-* de $L^\infty(\mu)$;*
   - *π es la única distribución cuasi-estacionaria del proceso $\hat{Y}$, la cual es equivalente a μ y tiene como parámetro asociado a θ;*
   - *f es una función acotada y θ-invariante y el proceso es θ-recurrente positivo.*

Demostración. La demostración de la segunda parte del teorema es igual a lo hecho para el Teorema 5.2 y por lo tanto la omitiremos. Para mostrar la primera parte notemos que las sucesiones $\{\theta_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ y $\{h_n\}_{n\in\mathbb{N}} \subseteq L^\beta(\mu_s)$ están acotadas, luego poseen subsucesiones convergentes a θ y débil-* convergentes a h.

Gracias al Teorema 3.3, para mostrar que la medida $hd\mu_s$ es θ-invariante basta deducir que para todo $w \in L^\infty(\hat{\mu}_s)$ se cumple

$$\langle \hat{P}_t(w), h\rangle_{\mu_s} \;=\; e^{-\theta t}\langle w, h\rangle_{\mu_s}.$$

Gracias a la Proposición 4.8, μ_s está bien concentrada y es 0-invariante, luego podemos utilizar el Teorema 4.2 para deducir que para todo $1 \leq k \leq \infty$ el operador $\hat{P}_t$ está bien definido desde $L^k(\mu_s)$ sobre sí mismo y que es continuo. Similarmente a lo hecho entonces para el Teorema 5.2 deducimos que

$$\begin{aligned}
\left|\langle \hat{P}_t(w), h\rangle_{\mu_s} - e^{-\theta t}\langle w, h\rangle_{\mu_s}\right| \;&\leq\; \left|\langle \hat{P}_t(w), h\rangle_{\mu_s} - \langle \hat{P}_t(w), h_n\rangle_{\mu_s}\right| + \left|\langle \hat{P}_t(w), h_n\rangle_{\mu_s} - \langle \hat{P}_t^n(w), h_n\rangle_{\mu_s}\right| \\
&\quad + \left|e^{-\theta_n t}\langle w, h_n\rangle_{\mu_s} e^{-\theta t}\langle w, h_n\rangle_{\mu_s}\right| + \left|e^{-\theta t}\langle h_n, w\rangle_{\mu_s} - e^{-\theta t}\langle w, h\rangle_{\mu_s}\right| \\
&\leq\; \left|\langle \hat{P}_t(w), h\rangle_{\mu_s} - \langle \hat{P}_t(w), h_n\rangle_{\mu_s}\right| + \left\|\hat{P}_t(w) - \hat{P}_t^n(w)\right\| \\
&\quad + \|w\|_\infty \left|e^{-\theta t} - e^{-\theta t}\right| + e^{-\theta t}\left|\langle w, h_n\rangle_{\mu_s} - \langle w, h\rangle_{\mu_s}\right|.
\end{aligned}$$

Notemos que como $w \in L^\infty(\mu_s)$, se tiene $P_t(w) \in L^\infty(\mu_s) \subseteq L^\alpha(\mu_s)$. Así, gracias a la convergencia débil-* de $\{h_n\}_{n\in\mathbb{N}}$ el primer y último término de la expresión anterior deben ser pequeños para valores grandes de n, y gracias a la Proposición 5.12 y la convergencia de los θ_n los términos restantes también serán pequeños. Se deduce que para todo $\varepsilon > 0$

$$\left|\langle \hat{P}_t(w), h\rangle_{\mu_s} - e^{-\theta t}\langle w, h\rangle_{\mu_s}\right| \;\leq\; \varepsilon,$$

de donde se deduce el resultado. □

Capítulo 6

Infinito como estado de entrada

En el presente capítulo resumiremos algunos de los principales resultados obtenidos para procesos de nacimiento y muerte en que el espacio de estados corresponde al conjunto de los números naturales. Para este caso estudiamos principalmente el trabajo realizado por Van Doorn en [30], donde se dan condiciones necesarias y suficientes sobre las tasas para la existencia y unicidad de una distribución cuasi-estacionaria. Es de particular interés el caso en que infinito es un estado de entrada, esto es, existen $t > 0$ y $m \in \mathbb{N}$ tales que

$$\inf_{n \geq m} \mathbb{P}_n(\mathcal{T}_m \leq t) > 0, \tag{6.1}$$

en que $\mathcal{T}_m$ es como en la Definición 2.19. Utilizando esta propiedad podremos mostrar que para ciertos procesos Y para los cuales infinito es un estado de entrada, sus aproximaciones $\{W^n\}_{n\in\mathbb{N}}$ poseen distribuciones cuasi-estacionarias, las cuales son únicas para cada W^n y convergen salvo subsucesión a una q.s.d. para Y. Mostraremos además que si para estos procesos existen vectores propios por la derecha $\{f_n\}_{n\in\mathbb{N}}$ los cuales son acotados, entonces existe un resultado análogo al Teorema 5.2 para Y.

6.1. Procesos de nacimiento y muerte en $\mathbb{N}$

En la presente sección enunciamos los principales resultados desarrollados en [30] para procesos de nacimiento y muerte en $\mathbb{N}$, los cuales conforman una de las familias más estudiadas de procesos de Markov absorbidos.

Definición 6.1 *En la presente sección introducimos un proceso de Markov X a tiempo continuo y cuyo espacio de estados son los números naturales. X corresponde a un proceso puro de saltos, en que la tasa de pasar desde n hasta $n+1$ es b_n y la tasa de pasar desde n hasta $n-1$ es λ_n. Asumiremos que $b_0 = \lambda_0 = 0$ y por lo tanto $\{0\}$ corresponde a un conjunto cementerio.*

El estudio de la existencia de distribuciones cuasi-estacionarias y de algunas propiedades de las mismas se basa principalmente en el estudio de las propiedades de expresiones construidas a partir de las tasas de transición, las cuales introducimos de la siguiente forma.

Definición 6.2 *Dadas las tasas de transición $\{b_n\}_{n\in\mathbb{N}}$ y $\{\lambda_n\}_{n\in\mathbb{N}}$ se definen:*

1. *$\rho_1 := 1$ y para cada $n \geq 2$, $\rho_{n+1} := \rho_n(b_n/\lambda_{n+1})$;*
2. *la serie A definida como $A := \sum_{n=1}^{\infty}(b_n\rho_n)^{-1}$;*
3. *la serie S definida como $S := \sum_{n=1}^{\infty}(b_n\rho_n)^{-1}\sum_{i=n+1}^{\infty}\rho_i$.*

En [14] se demuestra que la condición $A = \infty$ es equivalente a la absorción casi-segura del proceso en 0, y en [16], que las funciones $p_{ij}(t)$ definidas como

$$p_{ij}(t) := \mathbb{P}_i(X_t = j),$$

corresponden a las únicas soluciones de las ecuaciones

$$\begin{aligned} p'_{ij}(t) &= \lambda_i p_{i-1,j}(t) - (\lambda_i + b_i)p_{ij}(t) + b_i p_{i+1,j}(t) \qquad &\text{si} \quad i \geq 1,\ j \geq 0, \\ p'_{0,j}(t) &= 0 &\text{si} \quad j \geq 0. \end{aligned}$$

Karlin y McGregor mostraron en [15] que las funciones $p_{ij}(t)$ pueden ser representadas como

$$p_{ij}(t) = \rho_j \int_0^{\infty} e^{-xt} q_i(x) q_j(x) d\psi(x),$$

en que las funciones $\{q_i\}_{i\in\mathbb{N}}$ corresponden a una familia de polinomios definidos recurrentemente de la forma

$$\begin{aligned} b_n q_{n+1}(x) &:= (b_n + \lambda_n - x)q_n(x) - \lambda_n q_{n-1}(x), \\ b_1 q_2(x) &:= b_1 + \lambda_1 - x, \end{aligned}$$

y donde $q_1(x) \equiv 1$. Por otro lado, ψ es la única medida de probabilidad en $[0, \infty)$ para la cual los polinomios anteriores son ortogonales entre sí. Es directo notar que para todo x, la secuencia $\{q_n(x)\}_{n\in\mathbb{N}}$ corresponde a un vector propio por la derecha del generador infinitesimal del proceso, con valor propio asociado x. El siguiente resultado, utilizado por Van Doorn en [29] permite deducir para qué valores de x tales vectores propios son no negativos.

Proposición 6.1 *[29] Para cada $n \in \mathbb{N}$ el polinomio q_n posee exactamente n raíces distintas y positivas, las cuales notamos como $x_{n,1}, x_{n,2}, \ldots, x_{n,n}$. Estas raíces cumplen además que para todo $i, n \in \mathbb{N}$ con $i \leq n$*

$$x_{n+1,i} < x_{n,i} < x_{n+1,i+1}.$$

Es directo entonces deducir que para cada $i \in \mathbb{N}$, la secuencia $\{x_{n,i}\}_{n\geq i}$ es decreciente y acotada inferiormente por 0, luego converge a algún valor $\xi_i \geq 0$. De particular interés resulta entonces la equivalencia

$$x \leq \xi_i \iff \forall n \in \mathbb{N},\ q_n(x) > 0,$$

pues entrega condiciones necesarias y suficientes para la no negatividad de tales vectores propios.

El siguiente teorema muestra que la convergencia o divergencia de la serie S es determinante respecto a la cantidad de distribuciones cuasi-estacionarias del proceso.

Teorema 6.1 *[30, Teorema 3.2] Sea S como en la definición 6.2. Se tiene:*

- *si S diverge, entonces $\xi_1 = 0$ y no existe una q.s.d. para X o $\xi_1 > 0$ y para cada $0 < x \leq \xi_1$ existe una q.s.d. asociada al parámetro x;*
- *si S converge, entonces $\xi_1 > 0$ y existe una única q.s.d., la cual tiene como tasa de extinción a ξ_1.*

De esta forma, el estudio de la serie S permite revelar si la q.s.d. es única o no, y la razón de ello es que su convergencia o divergencia indica si infinito es o no un estado de entrada, como indica el siguiente resultado, que adaptamos desde el contexto de las difusiones.

Proposición 6.2 *[3, Proposiciones 7.6 y 7.10] Supongamos que la serie A diverge, entonces son equivalentes;*

- *existen $t > 0$ y $m \in \mathbb{N}$ tales que $\inf_{n \geq m} \mathbb{P}_n(\mathcal{T}_m \leq t) > 0$;*
- *$S < \infty$;*
- *$\sup_{k \geq 0} \mathbb{E}_k(\mathcal{T}_0) < \infty$;*
- *para todo $M \in \mathbb{R}^+$ existe $N \in \mathbb{N}$ tal que $\sup_{n \geq N} \mathbb{E}_n(e^{M\mathcal{T}_N}) < \infty$.*

Una propiedad muy importante y que no queda explícita en la proposición anterior es que la secuencia $\{\sup_{n \geq m} \mathbb{E}_n(\mathcal{T}_m)\}_{m \in \mathbb{N}}$ tiende a cero cuando m tiende a infinito. Esto es directo de la convergencia de S y de

$$\sup_{n \geq m} \mathbb{E}_n(\mathcal{T}_m) \;=\; \sum_{n=m+1}^{\infty} (b_n \rho_n)^{-1} \sum_{i=n+1}^{\infty} \rho_i,$$

cuya demostración podemos encontrar en [23]. En particular, utilizando la desigualdad de Markov podemos deducir el siguiente resultado.

Proposición 6.3 *Para todo $t > 0$ y $m \in \mathbb{N}$ se cumple $\inf_{n \geq m} \mathbb{P}_n(\mathcal{T}_m \leq t) > 0$.*

Demostración. Sea $t > 0$ fijo y tomemos $k \in \mathbb{N}$ tal que $2 \sup_{n \geq k} \mathbb{E}_n(\mathcal{T}_k) \leq t(1 - \varepsilon)$ en que $0 < \varepsilon < 1$ es cualquiera. Utilizando la desigualdad de Markov se deduce que para todo $n \geq k$ se tiene

$$\varepsilon \;\leq\; \mathbb{P}_n\left(\mathcal{T}_k \leq \frac{t}{2}\right).$$

Si $k \leq m$, se tiene $\mathcal{T}_m \leq \mathcal{T}_k$ y entonces se deduce el resultado. Si $k > m$, entonces para todo $n \geq k$

$$\mathbb{P}_n(\mathcal{T}_m \leq t) \;\geq\; \int_0^{t/2} \mathbb{P}_k(\mathcal{T}_m \leq t - s)\mathbb{P}_n(\mathcal{T}_k \in ds) \;\geq\; \mathbb{P}_k\left(\mathcal{T}_m < \frac{t}{2}\right)\varepsilon > 0,$$

y como para cada $m \leq n < k$ se tiene $\mathbb{P}_n(\mathcal{T}_m \leq t) > 0$ se deduce el resultado.

□

6.2. Distribuciones cuasi-estacionarias para W^n

En el capítulo anterior, a partir de las tasas de transición b_n, m_n, λ_n y g_n, fueron construidos los procesos W^n a valores en $\mathcal{A}^{-0}$, pero que se comportan como procesos a espacio de estados discreto al considerar particiones apropiadas de tal conjunto. El objetivo de esta sección es mostrar que bajo la hipótesis adicional de que infinito sea un estado de entrada para Y, es posible encontrar un resultado similar al Teorema 5.1 para cada W^n.

Comencemos notando que para el generador infinitesimal y el kernel de tasas de salto de W^n se tienen expresiones similares a (5.1), en que para $f : \mathcal{A} \to \mathbb{R}$ con $f(0) = 0$ localmente acotada, $\eta \in \mathcal{A}^{-0}$ y $B \subseteq \mathcal{A}$ se tiene

$$\begin{aligned} G_n f(\eta) = & \sum_{j=1}^{k} b(w_{D_j}, \zeta_H)[f(\eta + \delta_{y_j}) - f(\eta)] + \sum_{j=1}^{k} \lambda(w_{D_j}, \zeta_H)[f(\eta - \delta_{y_j}) - f(\eta)] \\ & + \sum_{j=1}^{k} m(w_{D_j}, \eta_H) \int_{\mathbb{T}} [f(\eta + \delta_z) - f(\eta)] g(w_{D_j}, z) d\sigma(z), \end{aligned}$$

$$Q_n(\eta, B) := \sum_{\substack{1 \le j \le k \\ \eta + \delta_{y_j} \in B}} b(w_{D_j}, \zeta_H) \; + \sum_{\substack{1 \le j \le k \\ \eta - \delta_{y_j} \in B}} \lambda(w_{D_j}, \zeta_H) \; + \; \sum_{j=1}^{k} m(w_{D_j}, \zeta_H) \int\limits_{\eta + \delta_z \in B} g(w_{D_j}, z) d\sigma(z),$$

donde $\vec{\eta} = (y_1, y_2, \ldots, y_k)$ y las notaciones w_D y ζ_H son como en el capítulo anterior. A pesar de que las tasas de transición de W^n no son continuas, éstas son localmente acotadas, luego los resultados del primer y tercer capítulo se mantienen válidos salvo aquellos relacionados con la existencia de q.s.d.. Para garantizar la existencia de tales distribuciones asumiremos una hipótesis adicional sobre las tasas de transición de Y.

Hipótesis 4 (HI) *Existen secuencias $\{b'_k\}_{k\in\mathbb{N}}$ y $\{\lambda'_k\}_{k\in\mathbb{N}}$ en $\mathbb{R}^+$ tales que $\forall \eta \in \mathcal{A}^{-0},\ y \in \{\eta\}$,*

$$\|\eta\|(b_y(\eta) + m_y(\eta)) \le b'_{\|\eta\|} \qquad y \qquad \|\eta\|\lambda_y(\eta) \ge \lambda'_{\|\eta\|},$$

y que, si son utilizadas para construir S y $\{\rho_j\}_{j\in\mathbb{N}}$ como en la definición 6.2, entonces $S < \infty$ y $b'_n \rho_n \xrightarrow[n\to\infty]{} 0$.

Notemos que gracias a la hipótesis (H1) se tiene

$$\sup_{\eta \in \mathcal{A}^{-0}} \sup_{y \in \{\eta\}} (b_y(\eta) + m_y(\eta)) \; = \; C < \infty,$$

luego, para que se cumpla (HI) en el caso de Y, basta tomar $b'_k := Ck$ y tasas de muerte λ'_k adecuadas. Un cálculo sencillo pero algo extenso muestra que es suficiente que las tasas de muerte tengan un crecimiento polinomial más rápido que lineal, es decir, $\lambda'_k := ck^{\beta}$ en que $\beta > 1$. En tal caso la serie S es convergente y además

$$b'_n \pi'_n \; = \; C \left(\frac{C}{c}\right)^{n-1} (n!)^{1-\beta} \; \xrightarrow[n\to\infty]{} 0,$$

luego en este caso se cumple (HI).

En lo que resta del capítulo asumiremos que Y cumple las hipótesis (H) junto con (HI) luego por construcción de los procesos W^n, éstos cumplen las mismas hipótesis. El siguiente resultado muestra que la hipótesis introducida en este capítulo permite acotar tanto $\|Y\|$ como los procesos $\|W^n\|$ por un proceso de nacimiento y muerte en $\mathbb{N}$ para el cual infinito es un estado de entrada.

Proposición 6.4 *Supongamos que Y cumple (HI) y sea X un proceso de nacimiento y muerte en $\mathbb{N}$ con tasas $\{b'_k\}_{k\in\mathbb{N}}$ y $\{\lambda'_k\}_{k\in\mathbb{N}}$ como en la hipótesis. En tal caso tanto $\|Y\|$ como los procesos $\|W^n\|$ están dominados estocásticamente por X.*

Demostración. Realizaremos la demostración de esta proposición únicamente para Y pues para cada W^n la demostración es idéntica. Consideremos un coupling F para los procesos Y y X tomando valores en el conjunto

$$\mathcal{G} \ := \ \{(\eta, k) \in \mathcal{A} \times \mathbb{N}, \ \|\eta\| \leq k\}.$$

El proceso F es un proceso de Markov de saltos cuyo kernel de tasas de transición $J((\eta, m), B)$ está definido como

$$\begin{array}{rcll}
J((\eta,k), A \times \{k\}) & = & \mathbf{1}_{\{k>\|\eta\|\}} Q(\eta, A) & \text{si } A \subseteq \mathcal{A}_{\|\eta\|+1}; \\
J((\eta,k), A \times \{k+1\}) & = & \mathbf{1}_{\{k=\|\eta\|\}} Q(\eta, A) & \text{si } A \subseteq \mathcal{A}_{\|\eta\|+1}; \\
J((\eta,k), \{\eta\} \times \{k+1\}) & = & b'_k - \mathbf{1}_{\{k=\|\eta\|\}} Q(\eta, \mathcal{A}_{\|\eta\|+1}); & \\
J((\eta,k), \{\eta - \delta_y\} \times \{k\}) & = & \eta_y \lambda_y(\eta) - \mathbf{1}_{\{k=\|\eta\|\}} \frac{\eta_y}{k} \lambda'_k & \text{si } y \in \{\eta\}; \\
J((\eta,k), \{\eta - \delta_y\} \times \{k-1\}) & = & \mathbf{1}_{\{k=\|\eta\|\}} \frac{\eta_y}{k} \lambda'_k & \text{si } y \in \{\eta\}; \\
J((\eta,k), \{\eta\} \times \{k-1\}) & = & \mathbf{1}_{\{\|\eta\|<k\}} \lambda'_k; & \\
J((0,m), \{0\} \times \{m+1\}) & = & b'_k; & \\
J((0,m), \{0\} \times \{m-1\}) & = & \lambda'_k. &
\end{array}$$

Por hipótesis cada una de estas tasas es no negativa y por lo tanto el proceso está bien definido. Notemos que para este proceso se tiene

$$J((\eta, \|\eta\|), \mathcal{A}_{>\|\eta\|} \times \{\|\eta\|\}) = 0,$$

y por lo tanto nunca sale de $\mathcal{G}$. Notando que los kernel de transición marginales corresponden a los de Y y X se deduce la dominación.

$\square$

Gracias a la proposición anterior, para toda función $f : \mathbb{N} \to \mathbb{R}$ creciente se tiene

$$\mathbb{E}_\eta(f(\|Y_t\|)) \ \leq \ \mathbb{E}_{\|\eta\|}(f(X_t)), \tag{6.2}$$

y por otro lado, como $\|Y\|$ está dominado por X, la variable $\mathcal{T}_k^Y$ está dominada estocásticamente por $\mathcal{T}_k^X$, en que ambas variables se definen como en la Definición 2.19. Esto es, para cada función $f : \mathbb{R} \to \mathbb{R}$ creciente se tiene

$$\mathbb{E}_\eta(f(\mathcal{T}_k^Y)) \ \leq \ \mathbb{E}_{\|\eta\|}(f(\mathcal{T}_k^X)). \tag{6.3}$$

Para poder encontrar un resultado similar al Teorema 5.1 observamos primero que se tiene la siguiente proposición cuya demostración omitiremos pues es idéntica a lo hecho para la Proposición 5.9.

Proposición 6.5 *Sea $n \in \mathbb{N}$ cualquiera, entonces para todo par de conjuntos $H, H' \in \mathcal{H}_n$, todo par de configuraciones $\eta_1, \eta_2 \in H$ y todo $t \geq 0$ se cumple*

$$\mathbb{P}_{\eta_1}(\hat{W}_t^n \in H') = \mathbb{P}_{\eta_2}(\hat{W}_t^n \in H').$$

Utilizaremos la proposición anterior para probar el siguiente resultado.

Teorema 6.2 *[21, Teorema 1] Sea $n \in \mathbb{N}$ cualquiera y W^n definido como antes. Para este proceso existe una única q.s.d. π_n y además existe $0 < \gamma < 1$ independiente de n tal que para toda medida de probabilidad ν sobre $\mathcal{A}^{-0}$ y $t \geq 1$ se cumple*

$$\|\mathbb{P}_\nu(W_t^n \in \cdot \mid t < \mathcal{T}) - \pi_n\|_{TV} \ \leq \ 2\gamma^{t-1},$$

en que $\|\cdot\|_{TV}$ es la norma de variación total y $\mathcal{T}$ es el tiempo de absorción para W^n.

Demostración. Al igual que en el Teorema 5.1 mostraremos que existe una medida ρ finita y no idénticamente nula tal que para todo $\eta \in \mathcal{A}^{-0}$ y $s \geq 1$ se cumple

$$\mathbb{P}_\eta(W_1^n \in \cdot \mid s < \mathcal{T}) \ \geq \ \rho. \tag{6.4}$$

Para ello realizaremos sutiles variaciones a lo hecho en el Teorema 1 de [21], comenzando por mostrar las siguientes propiedades.

- Existe $\alpha > 0$ independiente de n tal que para todo $\eta \in \mathcal{A}^{-0}$ y $A \subseteq \mathcal{A}_1$ medible se tiene

$$\mathbb{P}_\eta(W_1^n \in A) \ \geq \ \alpha\mu(A). \tag{6.5}$$

En efecto, utilizando la propiedad de Markov fuerte se tiene

$$\begin{aligned} \mathbb{P}_\eta(W_1^n \in A) \ &\geq \ \mathbb{P}_\eta\left(W_1^n \in A \ \wedge \ \mathcal{T}_1 < \tfrac{1}{2}\right) \\ &\geq \ \int_0^{\frac{1}{2}} \int_{\mathcal{A}_1} \mathbb{P}_{\eta'}(W_{1-s}^n \in A)\mathbb{P}_\eta\left(\mathcal{T}_1 \in ds \ \wedge \ W_s^n \in d\eta'\right), \end{aligned}$$

pero gracias a la desigualdad (5.4) existe α' independiente de s, η' y n tal que

$$\mathbb{P}_{\eta'}(W_{1-s}^n \in A) \ \geq \ \alpha'\mu(A),$$

podemos deducir

$$\mathbb{P}_\eta(W_1^n \in A) \ \geq \ \mathbb{P}_\eta\left(\mathcal{T}_1 \leq \frac{1}{2}\right)\alpha'\mu(A).$$

Ahora, tomando $f = \mathbf{1}_{t>1/2}$ en la ecuación (6.3) se tiene

$$\mathbb{P}_{\|\eta\|}(\mathcal{T}_1^X \leq 1/2) \ \leq \ \mathbb{P}_\eta(\mathcal{T}_1 \leq 1/2),$$

donde el primer término está acotado inferiormente por algún $\beta > 0$ independiente de $\|\eta\|$ y de n gracias a la Proposición 6.3. Tomando $\alpha = \beta\alpha'$ se deduce (6.4).

- Existe $\kappa > 0$ independiente de t y de n tal que

$$\kappa \sup_{\eta' \in \mathcal{A}^{-0}} \mathbb{P}_{\eta'}(\mathcal{T} > t) \;\leq\; \inf_{\eta \in \mathcal{A}_1} \mathbb{P}_{\eta}(\mathcal{T} > t). \tag{6.6}$$

Sea $\overline{Q} > 0$ una cota positiva para Q sobre $\mathcal{A}_1$, para la cual, gracias a las Proposiciones 6.2 y 6.4 existen $k \in \mathbb{N}$ y c_1 tales que

$$\sup_{\eta \in \mathcal{A}^{-0}} \mathbb{E}_{\eta}(e^{\overline{Q}\mathcal{T}_k}) \;\leq\; \sup_{n \in \mathbb{N}} \mathbb{E}_n(e^{\overline{Q}\mathcal{T}_k^X}) \;\leq c_1 \;<\; \infty, \tag{6.7}$$

y en que si además consideramos τ_1 el primer tiempo de salto de W^n entonces

$$\inf_{\eta \in \mathcal{A}_1} \mathbb{P}_{\eta}(W_t^n \in \mathcal{A}_{\leq k}) \;\geq\; \inf_{\eta \in \mathcal{A}_1} \mathbb{P}_{\eta}(\tau_1 > t) \;=\; e^{-\overline{Q}t}. \tag{6.8}$$

Buscamos acotar superiormente $\mathbb{P}_{\eta}(\mathcal{T} > t)$, dividiéndolo para ello como $\mathbb{P}_{\eta}(\mathcal{T}_k > t) + \mathbb{P}_{\eta}(\mathcal{T}_k \leq t < \mathcal{T})$, donde el primer término es fácil de acotar utilizando la desigualdad de Chernoff,

$$\mathbb{P}_{\eta}(\mathcal{T}_k > t) \;\leq\; e^{-\overline{Q}t}\mathbb{E}_{\eta}(e^{\overline{Q}\mathcal{T}_k}) \;\leq\; \inf_{\eta \in \mathcal{A}_1} \mathbb{P}_{\eta}(W_t^n \in \mathcal{A}_{\leq k})c_1 \;\leq\; \inf_{\eta \in \mathcal{A}_1} \mathbb{P}_{\eta}(\mathcal{T} > t)c_1.$$

Para el segundo término, gracias a la desigualdad (5.6) existe $\kappa' > 0$ independiente de t tal que

$$\inf_{\eta \in \mathcal{A}_{\leq k}} \mathbb{P}_{\eta}(\mathcal{T} > t) \;\geq\; \kappa' \sup_{\eta' \in \mathcal{A}_{\leq k}} \mathbb{P}_{\eta'}(\mathcal{T} > t),$$

y entonces, utilizando (6.7), (6.8) y la propiedad de Markov se tiene

$$\begin{aligned}
\mathbb{P}_{\eta}(\mathcal{T}_k \leq t < \mathcal{T}) \;&\leq\; \mathbb{E}_{\eta}(e^{\overline{Q}\mathcal{T}_k}) \sup_{\eta' \in \mathcal{A}_{\leq k}} \sup_{s \in [0,t]} e^{-\overline{Q}s}\mathbb{P}_{\eta'}(t - s < \mathcal{T}) \\
&\leq\; \tfrac{c_1}{\kappa'} \sup_{s \in [0,t]} \inf_{\eta' \in \mathcal{A}_1} \mathbb{P}_{\eta'}(W_s^n \in \mathcal{A}_{\leq k}) \inf_{\eta'' \in \mathcal{A}_{\leq k}} \mathbb{P}_{\eta''}(t - s < \mathcal{T}) \\
&=\; \tfrac{c_1}{\kappa'} \inf_{\eta' \in \mathcal{A}_1} \mathbb{P}_{\eta'}(t < \mathcal{T}),
\end{aligned}$$

luego se tiene

$$\mathbb{P}_{\eta}(\mathcal{T} > t) \;=\; \mathbb{P}_{\eta}(\mathcal{T}_k > t) + \mathbb{P}_{\eta}(\mathcal{T}_k \leq t < \mathcal{T}) \;\leq\; c_1 \left(1 + \frac{1}{\kappa'}\right) \inf_{\eta' \in \mathcal{A}_1} \mathbb{P}_{\eta'}(t < \mathcal{T}),$$

y tomando el supremo en $\mathcal{A}^{-0}$ se deduce el resultado.

Para encontrar la medida ρ que cumple (6.4) seguimos el mismo razonamiento utilizado en el Teorema 5.1, en que para $\eta \in \mathcal{A}^{-0}$, $A \subseteq \mathcal{A}_1$ medible y $s \geq 1$ se tiene

$$\begin{aligned}
\mathbb{P}_{\eta}(s < \mathcal{T})\mathbb{P}_{\eta}(W_1^n \in A \mid s < \mathcal{T}) \;&=\; \mathbb{P}_{\eta}(\hat{W}_1^n \in A,\; s < \mathcal{T}) \\
&\geq\; \kappa\, \mathbb{P}_{\eta}(\hat{W}_1^n \in A) \sup_{\eta' \in \mathcal{A}^{-0}} \mathbb{P}_{\eta''}(s - 1 < \mathcal{T}) \\
&\geq\; \kappa\alpha\mu(A)\mathbb{P}_{\eta}(s < \mathcal{T}),
\end{aligned}$$

luego, la medida $\rho \equiv \alpha\kappa\mu(\cdot \cap \mathcal{A}_1)$ cumple la propiedad deseada. El resto de la demostración es idéntico a lo hecho en el Teorema 5.1.

$\square$

Del resultado anterior se desprende la existencia y unicidad de una distribución cuasi-estacionaria para cada W^n, sin embargo nada podemos decir sobre la existencia de un vector propio por la derecha positivo y acotado, pues ni siquiera para el caso de un proceso de nacimiento y muerte en $\mathbb{N}$ tal existencia es clara. Sin embargo, dentro de tales procesos existen ciertos casos específicos para los cuales la existencia de tales vectores propios acotados es conocida, y que permiten pensar que esto ocurre bajo condiciones no muy restrictivas. Consideremos por ejemplo el caso en que las tasas de transición son de la forma $b_n := b_1 n$ en que $b_1 > 0$ y

$$\lambda_n \;:=\; \begin{cases} \dfrac{\theta\left[2n^\beta - 1\right] + b_1 n \left[1 - \left(\dfrac{n}{n+1}\right)^\beta\right]}{\left[\left(\dfrac{n}{n-1}\right)^\beta - 1\right]} & \text{si} \quad n \geq 2 \\ (1 - 2^{-\beta}) b_1 + \theta & \text{si} \quad n = 1, \end{cases}$$

donde $\theta, \beta > 0$. Un cálculo sencillo muestra que para n grande la función anterior puede aproximarse por $\frac{2\theta}{\beta} n^{\beta+1}$, y que para este proceso la función $\psi : \mathbb{N} \to \mathbb{R}^+$ definida como

$$\psi(n) \;:=\; 2 - n^{-\beta}$$

es un vector propio por la derecha asociado a $\theta > 0$ y es acotada.

6.3. Convergencia de los vectores propios de W^n

En la presente sección buscamos encontrar resultados similares a los de la Sección 5.4, según los cuales los vectores propios maximales de los procesos $\hat{W}^n$ convergen a los vectores propios maximales de $\hat{Y}$. Para ello necesitamos primero un resultado similar a la Proposición 5.11, el cual permita ver de qué forma los procesos W^n aproximan a Y.

Proposición 6.6 *Sean $\{\hat{P}^n_t\}_{t\geq 0}$ y $\{\hat{P}_t\}_{t\geq 0}$ los semigrupos de transición de los procesos W^n y Y respectivamente. Para todo $\varepsilon > 0$ y $m \in \mathbb{N}$ existe $N \in \mathbb{N}$ tal que para todo $n \geq N$ y $A \subseteq \mathcal{A}_{\leq p}$ medible se cumple*

$$\sup_{\eta \in \mathcal{A}_{\leq m}} |P^n_t(\mathbf{1}_A)(\eta) - P_t(\mathbf{1}_A)(\eta)| \leq \varepsilon.$$

Demostración. Comenzaremos mostrando para $t > 0$ y $\varepsilon > 0$ como en el enunciado existe un natural $k > m$ tal que para todo $n \in \mathbb{N}$,

$$\sup_{\eta \in \mathcal{A}_{\leq m}} \mathbb{P}_\eta(\mathcal{T}^n_k \leq t) \leq \varepsilon.$$

En efecto, utilizando la desigualdad (6.3), para todo $\eta \in \mathcal{A}_{\leq m}$ y $k \in \mathbb{N}$ se tiene

$$\mathbb{P}_\eta\left(\mathcal{T}^n_k \leq t\right) \;\leq\; \mathbb{P}_{\|\eta\|}\left(\mathcal{T}^X_k \leq t\right) \;\leq\; \mathbb{P}_m\left(\mathcal{T}^X_k < \infty\right),$$

pero el último término está acotado por la probabilidad de que existan tiempos de salto $\tau_{n_{m+1}}, \tau_{n_{m+2}}, \ldots, \tau_{n_k}$ en que para cada τ_{n_j}, X pasa desde $j-1$ al estado j. A su vez, esta probabilidad está acotada por la expresión

$$\frac{b'_m b'_{m+1} \cdots b'_{k-1}}{(b'_m + \lambda'_m)(b'_{m+1} + \lambda'_{m+1}) \cdots (b'_{k-1} + \lambda'_{k-1})} \;\leq\; \frac{b'_m b'_{m+1} \cdots b'_{k-1}}{\lambda'_m \lambda'_{m+1} \cdots \lambda'_{k-1}} \;=\; \frac{b'_{k-1}\rho_{k-1}}{b'_{m-2}\rho_{m-2}},$$

y como las tasas b'_n y λ'_n de X cumplen las condiciones de las hipótesis (HI), esta última expresión tiende a cero cuando $k \to \infty$. De esta forma, si tomamos k suficientemente grande

$$\sup_{\eta \in \mathcal{A}_{\leq m}} \mathbb{P}_\eta(\mathcal{T}_k^n \leq t) \;\leq\; \mathbb{P}_m\left(\mathcal{T}_k^X < \infty\right) \;\leq\; \varepsilon,$$

y la elección de k sólo depende de X, luego es válida para cada W^n y para Y.

Notemos ahora que para cada $\eta \in \mathcal{A}_{\leq m}$, la expresión $|P_t^n(\mathbf{1}_A)(\eta) - P_t(\mathbf{1}_A)(\eta)|$ es menor a

$$\left|\mathbb{P}_\eta(W_t^n \in A \;\wedge\; \mathcal{T}_k^n > t) - \mathbb{P}_\eta(Y_t \in A \;\wedge\; \mathcal{T}_k^Y > t)\right| + \mathbb{P}_\eta(\mathcal{T}_k^n \leq t) + \mathbb{P}_\eta(\mathcal{T}_k^Y \leq t),$$

luego, para $k \in \mathbb{N}$ grande e independiente de n los dos últimos términos son menores a ε. Por otro lado, en el evento $\mathcal{T}_k^n, \mathcal{T}_k^Y > t$, los procesos W^n e Y toman valores en $\mathcal{A}_{k-1}$, y por lo tanto podemos seguir la demostración realizada para la Proposición 5.11 para deducir la existencia de $N \in \mathbb{N}$ tal que para $n \geq N$ se cumple

$$\left|\mathbb{P}_\eta(W_t^n \in A \;\wedge\; \mathcal{T}_k^n > t) - \mathbb{P}_\eta(Y_t \in A \;\wedge\; \mathcal{T}_k^Y > t)\right| \;\leq\; \varepsilon.$$

Esta cota se obtiene independientemente de la configuración inicial η, luego se concluye el resultado.

□

El resultado anterior permite deducir el siguiente Corolario, cuya demostración omitiremos pues es idéntica a la del corolario 5.1.

Corolario 6.1 *Para todo $\varepsilon > 0$, $m \in \mathbb{N}$ y $C > 0$ existe $N \in \mathbb{N}$ tal que para todo $n \geq N$ y toda función $h : \mathcal{A}^{-0} \to \mathbb{R}^+$ a soporte compacto con $h \leq C$ se cumple*

$$\sup_{\eta \in \mathcal{A}_{\leq m}} |\hat{P}_t^n(h) - \hat{P}_t(h)| \leq \varepsilon.$$

Finalmente, estamos en condiciones de extender el Teorema 5.2 para procesos Y a valores en $\mathcal{A}^{-0}$ y que cumplen la hipótesis (HI).

Teorema 6.3 *Supongamos que Y cumple con (HI) y sea $\{\pi_n\}_{n\in\mathbb{N}}$ la secuencia de q.s.d. asociadas a los procesos $\{W^n\}_{n\in\mathbb{N}}$, cuya existencia está garantizada por el Teorema 6.2. Tomando $\{\theta_n\}_{n\in\mathbb{N}}$ la sucesión de tasas de extinción de tales q.s.d. se tiene;*

1. *Existe una distribución cuasi-estacionaria π para Y tal que $\pi_n \to \pi$ débilmente salvo subsucesión.*

2. *Supongamos que para cada W^n existe una función f_n acotada y θ_n-invariante. Si existe $u : \mathcal{A}_{\leq p} \to \mathbb{R}^+$, μ-integrable tal que*

$$\forall n \in \mathbb{N}, \qquad \int_{\mathcal{A}_{\leq p}} f_n(\eta) u(\eta) d\mu(\eta) \ = \ 1,$$

y si se cumple la condición

$$\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty,$$

entonces se tienen los siguientes resultados:

- *existen $\theta \in \mathbb{R}$, una medida de probabilidad π y una función f tales que $\theta_n \to \theta$, $\pi_n \to \pi$ débilmente y $f_n \to f$ en la topología débil-* de $L^\infty(\mu)$;*
- *π es la única distribución cuasi-estacionaria del proceso Y, la cual es equivalente a μ y tiene como parámetro asociado a θ;*
- *f es una función acotada y θ-invariante y el proceso es θ-recurrente positivo.*

Demostración. 1. Siendo los procesos construidos como en la Definición 2.11 generalizaciones de procesos de nacimiento y muerte en $\mathbb{N}$, los resultados encontrados para Y pueden ser aplicados de igual forma a X, pues este proceso cumple con las hipótesis en (H). En particular, gracias al Lema 3.3 se deduce que para todo $\gamma > 0$ existen $a_1, \beta > 0$ tales que

$$\forall m \in \mathbb{N}, \qquad \mathbb{E}_m(e^{a_1 X_1}, \mathcal{T}^X > 1) \ \leq \ \gamma e^{a_1 m} + \beta.$$

Tomando $\psi : \mathbb{N} \cup \{0\} \to \mathbb{R}$ como $\psi(m) = e^{a_1 m}$ sobre $\mathbb{N}$ y $\psi(0) = 0$ se tiene

$$\forall m \in \mathbb{N}, \qquad \mathbb{E}_m(\psi(X_1)) \ \leq \ \gamma\psi(m) + \beta,$$

y además ψ es creciente, luego, gracias a la Proposición 6.4, para cada $n \in \mathbb{N}$ se tiene

$$\forall \eta \in \mathcal{A}^{-0}, \quad \mathbb{E}_\eta(e^{a_1\|W_1^n\|}, \mathcal{T}^n > 1) \ = \ \mathbb{E}_\eta(\psi(\|W_1^n\|)) \ \leq \ \mathbb{E}_{\|\eta\|}(\psi(X_1)) \ \leq \ \gamma e^{a_1\|\eta\|} + \beta.$$

Por otro lado, tomando $\overline{Q}$ una cota superior para Q sobre $\mathcal{A}_1$ (y por lo tanto, de Q_n), el Lema 3.2 indica que para cada $n \in \mathbb{N}$

$$e^{-\overline{Q}} \ \leq \ P_t^n(\mathbf{1}_{\mathcal{A}^{-0}}),$$

desigualdad que utilizamos para obtener

$$\mathbb{E}_\nu(e^{a_1\|W_1^n\|} \mid \mathcal{T}^n > 1) \ \leq \ e^{\overline{Q}t}\left(\gamma\langle\nu, e^{a_1\|\cdot\|}\rangle + \beta\right), \tag{6.9}$$

donde ν es cualquier medida de probabilidad sobre $\mathcal{A}^{-0}$. En particular, si consideramos $\gamma < e^{-\overline{Q}}$ y $K \in \mathbb{R}^*$ tal que $\beta(e^{-\overline{Q}} - \gamma)^{-1} < K$, entonces para todo $n \in \mathbb{N}$ y toda medida de probabilidad ν,

$$\langle\nu, e^{a_1\|\cdot\|}\rangle \leq K \implies \mathbb{E}_\nu\left(e^{a_1\|W_1^n\|} \mid \mathcal{T}^n > 1\right) \leq K. \tag{6.10}$$

Notemos que utilizando la propiedad de Markov de W^n se tiene

$$\mathbb{E}_\nu\left(e^{a_1\|W_{m+1}^n 1\|} \mid \mathcal{T}^n > m+1\right) \ = \ \mathbb{E}_{\nu^*}\left(e^{a_1\|W_m^n\|} \mid \mathcal{T}^n > m\right),$$

donde $\nu^* := \mathbb{E}_\nu \left(W_1^n \in \cdot \;\middle|\; \mathcal{T}_0^{W^n} > 1 \right)$ es una medida de probabilidad, luego utilizando esta igualdad y (6.9) repetidamente, para todo $m \in \mathbb{N}$ se tiene

$$\langle \nu, e^{a_1 \|\cdot\|} \rangle \leq K \implies \mathbb{E}_\nu \left(e^{a_1 \|W_m^n\|} \;\middle|\; \mathcal{T}_0^{W^n} > m \right) \leq K, \tag{6.11}$$

implicancia que generaliza a (6.10). Sean finalmente $m, j \in \mathbb{N}$ cualesquiera, y tomemos π_n la distribución cuasi-estacionaria de W^n. Dada una medida de probabilidad ν tal que $\langle \nu, e^{a_1\|\cdot\|} \rangle \leq K$ y definiendo ν^n como $\nu^n = \mathbb{E}_\nu(W_m^n \in \cdot | \mathcal{T}^n > m)$ se cumple

$$\begin{aligned} \langle \pi_n, e^{a_1\|\cdot\|} \mathbf{1}_{\mathcal{A}_{\leq j}} \rangle &= \langle \pi_n - \nu^{*n}, e^{a_1\|\cdot\|} \mathbf{1}_{\mathcal{A}_{\leq j}} \rangle + \langle \nu^{*n}, e^{a_1\|\cdot\|} \mathbf{1}_{\mathcal{A}_{\leq j}} \rangle \\ &\leq e^{a_1 j} \|\pi_n - \mathbb{E}_\nu(W_m^n \in \cdot | \mathcal{T}^n > m)\|_{TV} + K, \end{aligned}$$

luego si m es suficientemente grande, utilizando el Teorema 6.2 se tiene

$$\langle \pi_n, e^{a_1\|\cdot\|} \mathbf{1}_{\mathcal{A}_{\leq j}} \rangle \leq K + 1,$$

y como $j \in \mathbb{N}$ es cualquiera, por teorema de convergencia monótona,

$$\langle \pi_n, e^{a_1\|\cdot\|} \rangle \;\leq\; K + 1. \tag{6.12}$$

Para probar entonces que existe una distribución cuasi-estacionaria, notemos que como $0 \leq \theta_n \leq \overline{Q}$, la sucesión $\{\theta_n\}_{n\in\mathbb{N}}$ es acotada y posee subsucesión convergente, digamos a un valor θ. Por otro lado, la sucesión $\{\pi_n\}_{n\in\mathbb{N}}$ es acotada vista como elementos en el dual de $\mathcal{C}_b(\mathcal{A}^{-0})$, introducido en la Definición 3.4. De esta forma $\{\pi_n\}_{n\in\mathbb{N}}$ posee una subsucesión débil-* convergente a un elemento $\pi \in \mathcal{C}_b(\mathcal{A}^{-0})^*$.

Como ya se había mencionado, $\mathcal{A}^{-0}$ es localmente compacto pero no compacto, y por lo tanto el dual del espacio anterior corresponde al espacio de las medidas finitamente aditivias y regulares sobre $\mathcal{A}^{-0}$,es decir, a priori π no corresponde a una medida σ-aditiva. Sin embargo, gracias a (6.12) es sencillo ver que para toda función $h : \mathcal{A}^{-0} \to \mathbb{R}^*$ con $h \leq e^{a_1\|\cdot\|}$ se cumple $\pi(h) \leq K + 2$, y podemos entonces utilizar [4, Lema 4.1] para concluir que π es una medida de probabilidad.

Utilizando el Teorema 3.3, para mostrar que π es una q.s.d. basta con ver que para todo $t \geq 0$ y $h : \mathcal{A}^{-0} \to \mathbb{R}$ continua y acotada se cumple

$$\langle \pi, P_t(h) \rangle \;=\; e^{-\theta t} \langle \pi, h \rangle.$$

Para ello, como π_n es q.s.d. para W^n se tiene

$$\begin{aligned} \left| \langle \pi, P_t(h) \rangle - e^{-\theta t} \langle \pi, h \rangle \right| &\leq |\langle \pi, P_t(h) \rangle - \langle \pi_n, P_t(h) \rangle| \;+\; |\langle \pi_n, P_t(h) \rangle - \langle \pi_n, P_t^n(h) \rangle| \\ &\quad + \left| e^{-\theta_n t} \langle \pi_n, h \rangle - e^{-\theta t} \langle \pi_n, h \rangle \right| \;+\; \left| e^{-\theta t} \langle \pi_n, h \rangle - e^{-\theta t} \langle \pi, h \rangle \right| \\ &\leq |\langle \pi, P_t(h) \rangle - \langle \pi_n, P_t(h) \rangle| \;+\; \sup_{\eta \in \mathcal{A}_{\leq m}} \left| \hat{P}_t(h)(\eta) - \hat{P}_t^n(h)(\eta) \right| \\ &\quad + 2\pi_n(\mathcal{A}_{>m}) \;+\; \|h\|_\infty \left| e^{-\theta_n t} - e^{-\theta t} \right| \;+\; e^{-\theta t} \left| \langle \pi_n, h \rangle - \langle \pi, h \rangle \right|, \end{aligned}$$

y utilizando los mismos argumentos que en el Teorema 5.2 se deduce que para $n \in \mathbb{N}$ grande, tanto el primer término como los últimos dos deben ser pequeños. Que el segundo término sea pequeño es consecuencia del Corolario 6.1, mientras que para el tercer término se tiene

$$\pi_n(\mathcal{A}_{>m}) \;\leq\; e^{-a_1 m}\langle \pi_n, e^{a_1\|\cdot\|}\mathbf{1}_{\mathcal{A}>m}\rangle \;\leq\; e^{-a_1 m}(K+1),$$

y tomando m grande, se deduce que para $\varepsilon > 0$ cualquiera

$$\left|\langle \pi, P_t(h)\rangle - e^{-\theta t}\langle \pi, h\rangle\right| \;\leq\; \varepsilon.$$

Como esto se cumple para todo $\varepsilon > 0$, π es una q.s.d..

2. La demostración de la segunda parte del teorema es análoga a la demostración realizada para el Teorema 5.2, y la única diferencia con lo hecho ahí consiste en demostrar que

$$\langle h, \hat{P}_t(f)\rangle_\pi \;=\; e^{-\theta t}\langle h, f\rangle_\pi$$

para toda función $h \in L^\infty(\pi)$ positiva y a soporte compacto, en lugar de demostrarlo para funciones $h \in L^1(\pi)$, pues de esta forma podemos utilizar el Corolario 6.1 en lugar del Corolario 5.1.

$\square$

Capítulo 7

Conclusiones y trabajo futuro

En el presente trabajo se determinó que bajo la condición de acotamiento uniforme para las funciones f_n, correspondientes a vectores propios para los procesos $\hat{W}^n$, existe convergencia de tales funciones a un propio por la derecha acotado de $\hat{Y}$, y que en tal caso se da una convergencia similiar para los vectores propios por la izquierda (Teorema 5.2). Además, para el caso en que Y vuelve rápidamente desde infinito se encontraron condiciones para tener el mismo resultado (Teorema 6.3). Utilizando estos teoremas se deduce un método con el cual es posible aproximar los vectores propios del proceso tal y como se hizo en la Sección 5.5.

El enfoque de buscar vectores propios por la derecha utilizado en este trabajo es poco utilizado en la literatura de distribuciones cuasi-estacionarias, siendo reemplazado por lo general por un análisis profundo de la estructura del proceso. De esta forma el trabajo realizado sienta un precedente, donde, si bien no se obtuvo resultados del todo determinantes para la existencia de la función f buscada, sí se realizó un análisis interesante sobre las propiedades del semigrupo en el Teorema 4.2, el cual puede extenderse a otros modelos permitiendo la búsqueda de funciones λ-invariantes acotadas.

En cuanto al trabajo futuro, el principal y más obvio es el de encontrar alguna cota uniforme para las funciones λ-invariantes. Resulta intuitivo pensar que para un espacio de estados compacto en que todas las tasas son acotadas, un vector propio por la derecha debiese ser acotado, sin embargo no se ha encontrado aún algún parámetro del proceso que permita acotar tales funciones. Una pista para la existencia de tal parámetro puede encontrarse en la convergencia de la serie S introducida en la Tefinición 6.2, puesto que pareciera ser que en los procesos de nacimiento y muerte en $\mathbb{N}$, la convergencia de tal serie equivale a que los vectores propios por la derecha sean acotados. Por otro lado, ya se vió en el Teorema 6.2 que la convergencia de S implica un comportamiento del proceso similar al caso truncado.

Un último problema abierto es el de la continuidad de la función f encontrada. En la totalidad de las simulaciones numéricas, las funciones f_n convergen puntualmente a una función continua, lo que, junto con la compacidad del espacio en el caso truncado, implicaría que f es acotada. Esta propiedad se presenta como una alternativa, sin embargo los resultados de existencia en el espacio de funciones involucran el concepto de equicontinuidad, que pareciera no ser compatible con la naturaleza del semigrupo de transición.

Anexo A

Demostraciones

En el presente capítulo demostraremos las Proposiciones 2.3, 3.3 y 4.3, que son de un carácter más técnico.

Proposición 2.3

1. La sucesión $\{(\bar{Y}_n, \tau_n)\}_{n\in\mathbb{N}}$ es cadena de Markov homogénea;
2. las variables $\Delta\tau_{k+1}\big|_{\bar{Y}_k}$ en que $\Delta\tau_{k+1} = \tau_{k+1} - \tau_k$, son independientes de a pares y tienen distribución exponencial con parámetro $Q(\bar{Y}_k)$;
3. los elementos condicionados $\bar{Y}_{n+1}\big|_{\bar{Y}_n}$ tienen como distribución $\frac{Q(\bar{Y}_n,\cdot)}{Q(\bar{Y}_n)}$;
4. para todo $n \in \mathbb{N}$ las variables $\bar{Y}_{n+1}\big|_{\bar{Y}_n}$ y $\Delta\tau_{n+1}\big|_{\bar{Y}_n}$ son independientes;
5. la sucesión $\{\bar{Y}_n\}_{n\in\mathbb{N}}$ es cadena de Markov homogénea.

Demostración. Sean $0 \leq s_1 \leq s_2$ y definamos los conjuntos B_{s_1,s_2} como

$$B_{s_1,s_2} := \{B \in \Sigma,\ \forall(t,j,y,\theta) \in B,\ s_1 < t \leq s_2\} \subseteq \Sigma,$$

que utilizaremos para la construcción de $\mathcal{F}_{s_1,s_2}$, la σ-álgebra generada por la familia deelementos aleatorios

$$\{Y_0\} \cup \{M_c(\cdot,B), M_m(\cdot,B), M_d(\cdot,B)\}_{B\in B_{s_1,s_2}}.$$

De la Definición 2.9 es inmediato ver que para $k \in \mathbb{N}$, $\eta_0, \eta_1, \eta_2, \ldots, \eta_k \in \mathcal{A}$ y $0 < t_1 < t_2 < \cdots < t_k$, el evento

$$F := \{\bar{Y}_0 = \eta_0,\ \bar{Y}_1 = \eta_1,\ \tau_1 = t_1,\ \bar{Y}_2 = \eta_2,\ \tau_2 = t_2,\ \ldots,\ \bar{Y}_k = \eta_k,\ \tau_k = t_k\}$$

es $\mathcal{F}_{0,t_k}$-medible, y en este evento las variable $\bar{Y}_{k+1}$, τ_{k+1} y v_x^k toman la forma

$$
\begin{aligned}
\tau_{k+1}|_F &= \text{ínf}\{t > t_k,\ \textstyle\sum_{x=c,m,d} M_x((t_k,t] \times E_x(\eta_k,t_k,t)) > 0\}; \\
v_x^k|_F &= \int\limits_{(t_k,\infty)\times E_x(\eta_k)} \mathbf{1}_{(t_k,\tau_{k+1}]}(t)\delta_{H^j(\eta_k)} M_x(dt,dj,dz,d\theta) \quad \text{Cuando } x = c, m; \\
v_m^k|_F &= \int\limits_{(t_k,\infty)\times E_m(\eta_k)} \mathbf{1}_{(t_k,\tau_{k+1}]}(t)\delta_z M_m(dt,dj,dz,d\theta); \\
\bar{Y}_{k+1}|_F &= \eta_k + v_c^k|_F + v_m^k|_F - v_d^k|_F,
\end{aligned}
$$

Que son $\mathcal{F}_{t_k,\infty}$-medibles, luego, independientes de $\mathcal{F}_{0,t_k}$ pues tales σ-álgebras están construidas a partir de medidas de Poisson independientes y sobre ventanas disjuntas. Podemos deducir entonces que dados $B \subseteq \mathcal{A}$ y $C \subseteq [0,\infty)$

$$
\mathbb{P}(\bar{Y}_{k+1} \in B \ \wedge \ \tau_{k+1} \in C \mid F) \ = \ \mathbb{P}(\bar{Y}_{k+1}|_F \in B \ \wedge \ \tau_{k+1}|_F \in C),
$$

lo que prueba 1). Para calcular la distribución conjunta de estas variables, sean F como antes, $y \in \{\eta_k\}$, $z' \notin \{\eta_k\}$ y $t_{k+1} > t_k$. Si consideramos D_c como

$$
D_c := \{(t,j,z,\theta) \in E \mid t \in [t_{k+1}, t_{k+1}+dt],\ j \leq \|\eta_{k+1}\|,\ 0 \leq \theta \leq b_{H^j(\eta_k)}(\eta_k)\},
$$

entonces es directo que el evento

$$
\left\{\bar{Y}_{k+1}|_F = \eta_k + \delta_y \ \wedge \ \bar{T}_{k+1}|_F \in [t_{k+1}, t_{k+1}+dt)\right\}
$$

es igual a

$$
\left\{M_c(D_c) = 1,\ M_c(E_c^{k,t_{k+1}} - D_c) = 0,\ M_m(E_m^{k,t_{k+1}}) = 0,\ M_d(E_d^{k,t_{k+1}}) = 0\right\}.
$$

Gracias a la independencia de las medidas y el que cada variable M_x sigue una distribución de Poisson, obtenemos

$$
\mathbb{P}\left(\bar{Y}_{k+1}|_F = \eta_k + \delta_y \ \wedge \ \tau_{k+1}|_F \in [t_{k+1}, t_{k+1}+dt)\right) = \mu_E(D_c) e^{-\mu_E(E_c^{k,t_{k+1}} \cup E_m^{k,t_{k+1}} \cup E_d^{k,t_{k+1}})},
$$

y es sencillo verificar que

$$
\begin{aligned}
\mu_E(E_c^{k,t_{k+1}} \cup E_m^{k,t_{k+1}} \cup E_d^{k,t_{k+1}}) &= Q(\eta_k)[t_{k+1} - t_k]; \\
\mu_E(D_c) &= (\eta_k)_y b_y(\eta_k) dt,
\end{aligned}
$$

luego, obtenemos

$$
\mathbb{P}\left(\bar{Y}_{k+1}|_F = \eta_k + \delta_y \ \wedge \ \tau_{k+1}|_F \in [t_{k+1}, t_{k+1}+dt)\right) = (\eta_k)_y b_y(\eta_k) e^{-Q(\eta)[t_{k+1}-t_k]} dt,
$$

y análogamente

$$
\begin{aligned}
\mathbb{P}\left(\bar{Y}_{k+1}|_F = \eta_k - \delta_y \ \wedge \ \tau_{k+1}|_F \in [t_{k+1}, t_{k+1}+dt)\right) &= (\eta_k)_y \lambda_y(\eta_k) e^{-Q(\eta)[t_{k+1}-t_k]} dt; \\
\mathbb{P}\left(\{\bar{Y}_{k+1}|_F - \eta_k\} \in dz' \ \wedge \ \tau_{k+1}|_F \in [t_{k+1}, t_{k+1}+dt)\right) &= (\eta_k)_y m_y(\eta_k) g(y,z') e^{-Q(\eta)[t_{k+1}-t_k]} dt d\sigma(z').
\end{aligned}
$$

Utilizando las igualdades anteriores deducimos finalmente que

$$\mathbb{P}(\bar{Y}_{k+1} \in B \ \wedge \ \tau_{k+1} \in C \mid F) \ = \ Q(\eta_k, B) \int_C e^{-Q(\eta_k)[t-t_k]} dt,$$

en que si tomamos la diferencia de los tiempos, se tiene

$$\mathbb{P}(\bar{Y}_{k+1} \in B \ \wedge \ \tau_{k+1} - \tau_k \in C \mid F) \ = \ Q(\eta_k, B) \int_C e^{-Q(\eta_k)t} dt,$$

que depende únicamente de η_k, y que corresponde a la multiplicación de las probabilidades condicionales. A partir de esta última igualdad se deducen fácilmente 2), 3), 4) y 5).

□

Proposición 3.3 Sean $n, m \in \mathbb{N}$ con $n > m$. Utilizando la notación $\mathcal{T}_m := \mathcal{T}_{CL_m(0)}$ se tiene que para todo $x \in CL_n(0)$,

$$\mathbb{P}_x(\mathcal{T}_{m+1} \leq \mathcal{T}_m) = 1.$$

Demostración. Mostraremos que el resultado se tiene para todo $n > m$ y $x \in CL_n(0)$ inductivamente en m. Para ello supongamos que $m = 1$ y tomemos $n > 2$ (el caso $n = 2$ es trivial). Supongamos que existe $x \in CL_n(0)$ tal que $\mathbb{P}_x(\mathcal{T}_2 > \mathcal{T}_1) > 0$, y notemos que el evento $\{\mathcal{T}_2 > \mathcal{T}_1\}$ es equivalente a

$$\{\exists k \in \mathbb{N}, \ \bar{X}_k \in CL_1(0) \ \wedge \ \forall 1 \leq j < k, \ \bar{X}_j \notin CL_2(0)\},$$

luego es unión numerable de eventos y $\mathbb{P}_x(\mathcal{T}_2 > \mathcal{T}_1) > 0$ implica que existe $k \in \mathbb{N}$ mínimo tal que

$$\mathbb{P}_x(\bar{X}_k \in CL_1(0) \ \wedge \ \forall 1 \leq j < k, \ \bar{X}_j \notin CL_2(0)) > 0.$$

Notemos que en particular existe un conjunto $A \subseteq \mathcal{X}$ con $A \cap CL_2(0) = \emptyset$ tal que

$$\mathbb{P}_x(\bar{X}_k \in CL_1(0) \ \wedge \ \bar{X}_{k-1} \in A) > 0,$$

pero utilizando la propiedad de Markov fuerte se tiene

$$\int_A \mathbb{P}_y(\bar{X}_1 \in CL_1(0)) \mathbb{P}_x(\bar{X}_{k-1} = dy) > 0,$$

luego en particular, sobre A se tiene $\mathbb{P}_y(\bar{X}_1 \in CL_1(0)) > 0$, y entonces gracias a la propiedad de Markov, cada $y \in A$ cumple

$$\mathbb{P}_y(\bar{X}_2 = 0) \ \geq \ \int_{CL_1(0)} \mathbb{P}_z(\bar{X}_1 = 0) \mathbb{P}_y(\bar{X}_1 = dy) > 0,$$

luego, $y \in CL_2(0)$ o $y = 0$, pero como $\mathbb{P}_y(\bar{X}_1 \in CL_1(0)) > 0$ podemos descartar esta última opción, de donde se tiene una contradicción.

El caso general es similar, pues suponiendo $\mathbb{P}_x(\mathcal{T}_{m+1} > \mathcal{T}_m) > 0$, existe $k \in \mathbb{N}$ mínimo tal que

$$\mathbb{P}_x(\bar{X}_k \in CL_m(0) \ \wedge \ \forall 1 \leq j < k, \ \bar{X}_j \notin CL_{m+1}(0)) > 0,$$

luego existe $A \subseteq \mathcal{A}$ con $A \cap CL_{m+1}(0) = \emptyset$ tal que

$$\mathbb{P}_x(\bar{X}_k \in CL_m(0) \ \wedge \ \bar{X}_{k-1} \in A) > 0,$$

y utilizando la propiedad de Markov, sobre A se tiene $\mathbb{P}_y(\bar{X}_1 \in CL_m(0)) > 0$, luego

$$\mathbb{P}_y(\bar{X}_{m+1} = 0) \ \geq \ \int_{CL_m(0)} \mathbb{P}_z(\bar{X}_m = 0)\mathbb{P}_y(\bar{X}_1 = dy) > 0.$$

Como $y \notin CL_{m+1}(0)$, necesariamente existe un $j \leq m$ tal que $y \in CL_j(0)$. Sea J el máximo de los j tales que $\mathbb{P}_x(\bar{X}_{k-1} \in A \cap CL_j(0)) > 0$, luego, si $j = m$ se viola la minimalidad de k, y si $j < m$, el resultado se tiene por hipótesis de inducción, pues

$$\mathbb{P}_x(\bar{X}_{k-1} \in CL_j(0) \ \wedge \ \forall 1 \leq i < k, \ \bar{X}_i \notin CL_{j+1}(0)) > 0,$$

de donde $\mathbb{P}_x(\mathcal{T}_{j+1} > \mathcal{T}_j) > 0$, lo que es una contradicción.

$\square$

Proposición 4.3 Para los operadores CL y $\{CL_n\}_{n\geq 0}$ se cumple;

1. sean $\{B_n\}_{n\geq 0}$ medibles, entonces $CL(\cup_{n=1}^{\infty} B_n) = \cup_{n=1}^{\infty} CL(B_n)$;
2. sean $\{B_n\}_{n\geq 0}$ medibles, entonces $CL_1(\cup_{n=1}^{\infty} B_n) \subseteq \cup_{n=1}^{\infty} CL_1(B_n)$;
3. para $A \subseteq B$ medibles, $CL(A) \subseteq CL(B)$;
4. para $n \geq 1$ y B medible, $CL_n(B) = CL_1^{(n)}(B) - \cup_{j=1}^{n-1} CL_1^j(B)$ en que $CL_1^{(n)}$ es la composición n veces del operador CL_1;
5. para A medible se cumple $CL(CL(A)) = CL(A)$.

Demostración.

1. $\boxed{\subseteq}$ Sea $\eta \in CL(\cup_{n=1}^{\infty} B_n)$, luego gracias a la Proposición 4.2 existe $t > 0$ tal que

$$0 \ < \ \mathbb{P}_\eta(Y_t \in \cup_{n=1}^{\infty} B_n) \ \leq \ \sum_{n=0}^{\infty} \mathbb{P}_\eta(Y_t \in B_n).$$

En particular existe $k \in \mathbb{N}$ tal que $0 < \mathbb{P}_\eta(Y_t \in B_k)$, y utilizando nuevamente la Proposición 4.2 se tiene que $\eta \in CL(B_k) \subseteq \cup_{n=1}^{\infty} CL(B_n)$.

$\boxed{\supseteq}$ Sea $\eta \in \cup_{n=1}^{\infty} CL(B_n)$, luego existe $k \in \mathbb{N}$ tal que $\eta \in CL(B_k)$, es decir $\exists t > 0$ con

$$0 < \mathbb{P}_\eta(Y_t \in B_k) \leq \mathbb{P}_\eta(Y_t \in \cup_{n=1}^{\infty} B_n),$$

y gracias a la Proposición 4.2 necesariamente $\eta \in CL(\cup_{n=1}^{\infty} B_n)$.

2. Tomemos $\eta \in CL_1(\cup_{n=1}^{\infty} B_n)$, luego $\mathbb{P}_\eta(\bar{Y}_1 \in \cup_{n=1}^{\infty} B_n) > 0$, y como la unión es numerable, existe $k \in \mathbb{N}$ tal que

$$\mathbb{P}_\eta(\bar{Y}_1 \in B_k) > 0.$$

De esta forma, $\eta \in CL_1(B_k) \subseteq \cup_{n=1}^{\infty} CL_1(B_n)$.

3. Sea $\eta \in CL(A)$, luego, gracias a la Proposición 4.2 existe $t > 0$ tal que $\mathbb{P}_\eta(Y_t \in A) > 0$, y como $A \subseteq B$, $\mathbb{P}_\eta(Y_t \in B) > 0$, luego de la Proposición 4.2 se concluye que $\eta \in CL(B)$.

4. Mostremos inductivamente que para todo $n \in \mathbb{N}$,
$$CL_1^{(n)}(B) = \{\eta \in \mathcal{A},\ \mathbb{P}_\eta(Y_n \in B) > 0\}.$$
En efecto, para el caso $n = 1$ el resultado es trivial, y para el caso $n+1$ se tiene $\boxed{\subseteq}$ dado $\eta \in CL_1^{(n+1)}(B) = CL_1(CL_1^{(n)}(B))$ se tiene
$$\mathbb{P}_\eta(\bar{Y}_1 \in CL_1^{(n)}(B)) > 0,$$
y por hipótesis de inducción y la propiedad de Markov,
$$\mathbb{P}_\eta(\bar{Y}_{n+1} \in B) \ \geq \ \int_{CL_1^{(n)}(B)} \mathbb{P}_{\eta'}(\bar{Y}_n \in B)\mathbb{P}_\eta(\bar{Y}_1 = d\eta') > 0.$$
$\boxed{\subseteq}$ dado η tal que $\mathbb{P}_\eta(Y_{n+1} \in B) > 0$, utilizando la propiedad de Markov y la hipótesis de inducción se tiene
$$\mathbb{P}_\eta(Y_{n+1} \in B) \ = \ \int_{CL_1^{(n)}(B)} \mathbb{P}_{\eta'}(\bar{Y}_n \in B)\mathbb{P}_\eta(\bar{Y}_1 = d\eta') > 0,$$
de donde $\mathbb{P}_\eta(\bar{Y}_1 \in CL_1^{(n)}(B)) > 0$, es decir, $\eta \in CL_1(CL_1^{(n)}(B))$.

5. $\boxed{\subseteq}$ Sea $\eta \in CL(CL(A))$, luego, gracias a la Proposición 4.2 existe $n \in \mathbb{N}$ tal que $\mathbb{P}_\eta(\bar{Y}_n \in CL(A)) > 0$, y por otro lado, para cada $\eta' \in CL(A)$ existe un $m \in \mathbb{N}$ tal que $\mathbb{P}_\eta(\bar{Y}_m \in A) > 0$. Como hay una cantidad numerable de tales m, en particular, existe m tal que
$$\mathbb{P}_\eta(\bar{Y}_n \in CL(A) \cap J_m) > 0,$$
donde J_m se compone de los $\eta \in CL(A)$ que llegan en m saltos a A. De esta forma,
$$\mathbb{P}_\eta(\bar{Y}_{n+m} \in A) \ \geq \ \int_{CL(A)} \mathbb{P}_{\eta'}(\bar{Y}_m \in A)\mathbb{P}_\eta(\bar{Y}_n = d\eta') > 0.$$
$\boxed{\supseteq}$ Notemos que $A \subseteq CL(A)$, en efecto, como el proceso es de saltos, para $\eta \in A$ se tiene $\mathbb{P}_\eta(Y_t \in A) \geq \mathbb{P}_\eta(\tau_1 > t) > 0$, y entonces, utilizando la propiedad 3) aquí demostrada se tiene el resultado.

$\square$

Bibliografía

[1] Patrick Billingsley. *Convergence of Probability Measures*, volume 493 of *Wiley Series in Probability and Statistics*. John Wiley & Sons, Inc., second edition, 2009.

[2] Laird Arnault Breyer. *Quasistationarity and Conditioned Markov Processes*. PhD thesis, The University of Queensland, June 1997.

[3] Patrick Cattiaux, Pierre Collet, Amaury Lambert, Servet Martínez, Sylvie Méléard, and Jaime San Martín. Quasi-stationary distributions and diffusion models in population dynamics. *The Annals of Probability*, 37(5), 2009.

[4] Pierre Collet, Servet Martínez, Sylvie Méléard, and Jaime San Martín. Quasi-stationary distributions for structured birth and death processes with mutations. *Probability Theory and Related Fields*, 151(1-2):191–231, October 2011.

[5] Pierre Collet, Servet Martínez, and Jaime San Martín. *Quasi-Stationary Distributions Markov Chains, Diffusions and Dynamical Systems*. Springer, 2013.

[6] Nelson Dunford and Jacob Schwartz. *Linear operators. Part I: General theory.* Interscience Publishers Inc, 1957.

[7] William Feller. *An Introduction to Probability Theory and Its Applications*, volume I. John Wiley & Sons, Inc., 1966.

[8] P. Ferrari, H. Kesten, S. Martínez, and P. Picco. Existence of quasi-stationary distributions. a renewal dynamical approach. *The Annals of Probability*, 23(2):501–521, 1995.

[9] Joaquín Fontbona. Procesos puntuales de poisson y aplicaciones. VI Escuela de Primavera DIM, 2006.

[10] Nicolas Fournier and Sylvie Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *The Annals of Applied Probability*, 14(4), 2004.

[11] Boris Harlamov. *Continuous Semi-Markov Processes*. John Wiley & Sons, Inc., 2008.

[12] T. E. Harris. The existence of stationary measures for certain markov processes. *Third Berkeley Symposium on Mathematical Statistics and Probability*, 2, 1956.

[13] Roger Horn, Johnson A., and Charles R. *Topics in Matrix Analysis*. Cambridge Uni-

versity Press, 1994.

[14] Samuel Karlin and James McGregor. The classification of birth and death processes. *Transactions of the American Mathematical Society*, 86(2), 1957.

[15] Samuel Karlin and James McGregor. The differential equations of birth-and-death processes, and the stieltjes moment problem. *Transactions of the American Mathematical Society*, 85(2), 1957.

[16] J. H. B. Kemperman. An analytical approach to the differential equations of the birth-and-death process. *Michigan Math. J.*, 9(4), 1962.

[17] Claudia Knoche. Stochastic integrals and stochastic differential equations with respect to compensated poisson random measures in infinite dimensional hilbert spaces. Fakultät für Mathematik Universität Bielefeld, 2003.

[18] H. Krieger. Exponential random variables and explosions. Harvey Mudd College, 2008.

[19] Thomas Kurtz and Stewart Ethier. *Markov Processes Characterization and Convergence*. John Wiley & Sons, Inc., 2005.

[20] Servet Martínez. Notes and a remark on quasi-stationary distributions. *Pyrenees International Workshop on Statistics, Probability and Operations Research*, 2008.

[21] Servet Martínez, Jaime San Martín, and Denis Villemonais. Existence and uniqueness of a quasi-stationary distribution for markov processes with fast return from infinity. Applied Probability Trust, 2013.

[22] Sylvie Méléard and Villemonais Denis. Quasi-stationary distributions and population processes quasi-stationary distributions and population processes quasi-stationary distributions and population processes quasi-stationary distributions and population processes. *Probability Surveys*, 9, 2012.

[23] Antonina Mitrofanova. Lecture 4: Absorbing probability and mean time to absorption in birth and death process. applications to moran model. NYU, department of Computer Science NYU, department of Computer Science NYU, department of Computer Science, 2007.

[24] Ingemar Nåsell. The quasi-stationary distribution of the closed endemic sis model. *Advances in Applied Probability*, 28(3), 1996.

[25] S. Orey. *Lecture Notes on Limit Theorems for Markov Chain Transition Probabilities*. Van Nostrand-Reinhold, London, 1971.

[26] Pekka Tuominen and Richard Tweedie. Exponential decay and ergodicity of general markov processes and their discrete skeletons. *Advances in Applied Probability*, 11(4):784–803, December 1979.

[27] Richard Tweedie. r-theory for markov chains on a general state space i: Solidarity

properties and r-recurrent chains. *The Annals of Probability*, 2(5):840–864, 1974.

[28] Richard Tweedie. r-theory for markov chains on a general state space ii: r-subinvariant measures for r-transient chains. *The Annals of Probability*, 2(5):865–878, 1974.

[29] Erik Van Doorn. Conditions for exponential ergodicity and bounds for the decay parameter of a birth-death process. *Advances in Applied Probability*, 17(3), 1985.

[30] Erik Van Doorn. Quasi-stationary distributions and convergence of quasi-stationarity of birth-death processes. *Advances in Applied Probability*, 23(4), December 1991.

[31] D. Vere-Jones. Geometric ergodicity in denumerable markov chains. *Quart. J. Math.*, 13, 1962.

[32] D. Vere-Jones. Ergodic properties of nonnegative matrices. *Pacific. J. Math*, 22, 1967.