CIMAT

Centro de Investigación en Matemáticas, A.C.

Convergencia Débil: Límite de Difusión del Proceso de Edades en un Proceso de Galton-Waston a Tiempo Continuo

T E S I S

Que para obtener el grado de

Maestría en Ciencias

con Orientación en

Probabilidad y estadística

P r e s e n t a

Ernesto Ramos López

Director de Tesis:

Dr. Antonio Murillo Salas

Guanajuato, Gto. Noviembre de 2014

Convergencia débil: límite de difusión del proceso de edades en un proceso de Galton-Watson a tiempo continuo

Ernesto Ramos López

2014

A Miriam Guadalupe Baez Hernández.

Agradecimientos

A mis padres, Genoveva y Laurencio, por su apoyo incondicional y ánimo para conseguir mis metas. Gracias por creer nuevamente en mi.

A mis hermanos, Cesar, Lisbia, Oscar e Isahí, que aunque no lo hayan hecho explícitamente, supieron alentarme.

Al Dr. Antonio Murillo Salas, por su paciencia, orientación y apoyo. Gracias por su preocupación, consejos y motivación, muy particular, de animarme a seguir adelante en el área académico.

A los Dres. Jose Alfredo López Mimbela y Arnaud Charles Leo Jégousse, porque destinaron parte de su tiempo para la revisión de este trabajo, haciéndome sugerencias y aportaciones que contribuyeron a la mejora del mismo.

A *los profesores del CIMAT* que fueron parte fundamental para mi formación académica, y en particular, a los Dres. Víctor Manuel Pérez Abreu Carrión y Miguel Nakamura Savoy por la oportunidad brindada para estudiar el posgrado.

A *mi generación*, por sus enseñanzas, por brindarme su confianza y amistad, y haber compartido actividades extra académicas. Gracias por haber hecho mas placentero y divertido mi estancia en la posgrado. Un especial agradecimiento a Tulio, Sharo, Rodrigo, Sandra y Adan porque a pesar del poco tiempo de convivencia me recordaron el espíritu de compañerismo.

A *la población mexicana*, porque hace posible la existencia de becas de posgrado, de las cuales la beca No. 347527 me fue otorgada vía el Consejo Nacional de Ciencia y Tecnología. Tambien al CIMAT por las becas dadas en mi estancia.

A la familia Baez Hernández, por su apoyo incodicional.

Finalmente, *a Miriam G. Baez Hernández*, por su apoyo, comprensión, paciencia y compañía incodicional. Gracias por haber aceptado ser parte de esta etapa de mi vida. Es poco el espacio en esta hoja para describir mi agradecimiento hacia ti.

Índice general

Introducción

El propósito de esta tesis es presentar en forma autocontenida una recopilación de resultados fundamentales en la teoría de convergencia de procesos estocásticos con valores en un espacio métrico. Además, se describirá la forma en que dicha teoría puede ser usada para obtener un límite de difusión del proceso de edades en un proceso de Galton-Watson a tiempo continuo siguiendo las ideas presentadas en [6].

La teoría de convergencia de procesos fue impulsado por los trabajos de Anatoliy Volodimirovich Skorokhod. El primer resultado que comenzó el desarrollo de esta teoría fue el principio de invarianza de Donsker (1953). Éste establece que una sucesión de procesos estocásticos, construida a partir de una familia de variables aleatorias independientes e identicamente distribuidas con tayectorias en $C([0,1])$ (el espacio de funciones real valuadas en [0,1] que son continuas y acotada), converge débilmente al proceso estocástico con trayectorias en $C([0,1])$ conocido como proceso de Wiener. Más tarde (1956), Yu. V. Prohorov estableció el teorema general sustituyendo el espacio $C([0,1])$ por un espacio métrico completo y separable. Dada una sucesión de procesos aleatorios relativamente compacta con trayectorias en un espacio métrico completo y separable, y cuyas distribuciones finito dimensionales convergen débilmente a un proceso dado, la sucesión de procesos estocásticos convergen débilmente al proceso dado. Este nuevo resultado dió lugar al problema de seleccionar una métrica adecuada en E y determinar criterios de compacidad relativa para una sucesión de procesos estocásticos.

Al mismo tiempo, comenzaron a aparecer procesos cuyas trayectorias no necesariamente eran continuas, lo que provocó considerar el espacio de funciones definidas en $[0,1]$ que son continuas por la derecha y con límites por la izquierda, comunmente denotado por $D([0,1])$. Con el fin de usar el teorema de Prohorov en situaciones donde los procesos tuvieran trayectorias en $D([0,1])$, el espacio debía ser metrizado ya que la métrica uniforme definida en $C([0,1])$ no era apropiada en dicho espacio. Este problema fue resuelto por Skorokhod (1956) en [23], quien, posteriormente, aplicó sus resultados a procesos particulares, a saber, convergencia de procesos con incrementos independientes y procesos de Markov. La teoría fue más tarde extendida a $\mathcal{D}([0,\infty),E)$, el espacio de funciones en $[0,\infty)$ y valores en un espacio métrico E que son continuas por la derecha y tienen límites por la izquierda, por C.J. Stone (1963) y T. Lindvall (1973). Prohorov (1961) promovió el estudio de esta teoría considerando procesos estocásticos con valores en un espacio de medidas, la cual prosiguió su desarrollo con H. Debes (1970), T.E. Harris (1971), P. Jagers (1974), Olav Kallenberg (1973,1986,1996) y Thomas G. Kurtz(1986).

La contribución de este trabajo radica en recopilar los resultados fundamentales de algunos textos clásicos tales como [4, 8, 22], en el cual se estudian convergencia débil de medidas de probabilidad definidas en un espacio métrico y el espacio $\mathcal{D}([0,\infty),E)$ con una métrica que induce la topología de Skorokhod. Además, se completan y revisan varios argumentos en las demostraciones de los resultados más importantes de la teoría de convergencia de procesos estocásticos con trayectorias en $\mathcal{D}([0,\infty),E)$. Cabe mencionar, que no se pretende ninguna originalidad más allá de presentar en forma integrada los resultados fundamentales de esta útil teoría, y la forma en que ésta sirve para probar convergencia de procesos.

La tesis está divididad en tres capítulos. En el primer capítulo se presentan, sin demostración, algunos resultados referentes a la teoría de ramificación, el problema de la martingala y a la teoría de medidas aleatorias. Se expone la transformada de Laplace del proceso de ramificación con espacio de estados continuo en el caso crítico o subcrítico, se plantea el problema de la martingala, cuya importancia radica en caracterizar a los process de Markov como una solución al problema de la martingala y se introduce el espacio de medidas de Radon.

En el segundo capítulo, se desarrollan las herramientas suficientes para establecer el *teorema principal* (Teorema 2.2.20), que afirma que una sucesión de procesos estocásticos con trayectorias en $\mathcal{D}([0,\infty),E)$ tensa y cuyas distribuciones finito dimensionales convergen débilmente a un proceso dado, entonces la sucesión converge débilmente al proceso dado. Se definen la convergencia débil y la tensión de una sucesión de medidas de probabilidad definidas en un espacio métrico. Se presenta el célebre Teorema de Prokhorov (1956), el cual relaciona la propiedad de tensión con la propiedad de compacidad relativa de una familia de medidas de probabilidad. Posteriormente, se analiza el espacio métrico $\mathcal{D}([0,\infty),E)$, el cual es un espacio separable y completo bajo una métrica adecuada, y se estudian propiedades fundamentales de este espacio. Este espacio resulta de gran interés ya que muchos procesos estocásticos tienen trayectorias con saltos, es decir, procesos cuyas trayectorias pertenecen a $\mathcal{D}([0,\infty),E)$, como, por ejemplo, el proceso de Poisson. Finalmente, se provee un criterio de tensión de procesos estocásticos con valores en el espacio de medidas de Radon.

Por último, en el tercer capítulo, se aplicará el Teorema 2.2.20 a una sucesión de procesos relacionadas con de copias independientes del proceso de edades. Basicamente, el proceso de edades estudia la variación en el tiempo del tamaño de una población, cuyos *miembros están caracterizado por su edad* y considerando las modificaciones producidas por los fenómenos de nacimiento, muerte y envejecimiento de los miembros de la población. Como consecuencia, se obtendrá un límite de difusión en el sentido de convergencia débil.

Notación

Símbolo	**Significado**
A'	El complemento del conjunto A.
$\mathbb{N}$	$\{1,2,3,\ldots\}$.
$\mathbb{R}$	Números reales.
$\mathbb{R}^+$	Números reales no negativos.
$\mathbb{R}^d$	Espacio euclideano $d-$dimensional, $d \in \mathbb{N}$.
$\mathbb{R}^\infty$	Espacio de sucesiones $x = (x_1, x_2, \ldots)$ de número reales.
càdlàg	Funciones que son continuas por la derecha con límites por la izquierda.
$C([0,1])$	El espacio de funciones real valuadas continuas definidas en $[0,1]$.
$D([0,1])$	El espacio de funciones *càdlàg* definidas en $[0,1]$.
$\mathcal{C}_b(E)$	La familia de funciones $f : E \to \mathbb{R}$ continuos y acotadas.
$\mathcal{C}_K(E)$	La familia de funciones $f : E \to \mathbb{R}$ continuos, acotadas y con soporte compacto.
$\mathcal{C}_b^1(E)$	La familia de funciones $f : \mathbb{R}^+ \to \mathbb{R}$ con primera derivada continua y acotada.
$\mathcal{C}_{b,+}^1(E)$	El subconjunto de elementos no negativos con soporte compacto de $\mathcal{C}_b^1(E)$.
$M(E)$	La familia de funciones $f : E \to \mathbb{R}$ continuos, acotadas y medibles.
$\lVert\cdot\rVert$	Norma del supremo definido en $C([0,1])$.
$\lVert\cdot\rVert_\infty$	Norma del supremo definido en $\mathcal{C}_b(\mathbb{R})$.
$\mathcal{M}(E)$	El espacio de medidas en E que son Borel-Radon.
$\mathcal{M}_b^+(\mathbb{R}^+)$	El espacio de las medidas de Borel-Radon no negativas en $\mathbb{R}^+$ que son finitas.
$\mathbf{N} = \{N_t, t \geq 0\}$	Proceso de nacimiento y muerte de parámetro λ.
$\mathbf{X} = \{\mathbf{X}_t, t \geq 0\}$	El proceso de edades.
$T_t\phi(x)$	$\mathbf{E}_x\left[\langle X_t, \phi\rangle\right]$ para $t \geq 0$, $\phi \in \mathcal{C}_b^+(\mathbb{R}^+)$.
$K_t\phi(x)$	$-\log \mathbf{E}_x\left[\exp\left\{-\langle \mathbf{X}_t, \phi\rangle\right\}\right]$, $\phi \in \mathcal{C}_b^+(\mathbb{R}^+)$.

Capítulo 1

Preliminares

En este capítulo, se presentan definiciones y resultados que se usarán en capítulos posteriores. Se expone de manera general, y sin demostración, los resultados referente a la forma general de la transformada de Laplace de los procesos de ramificación con espacio de estado continuo en el caso crítico o subcrítico (con especial atención a la transformada de Laplace de la difusión de Feller) y el problema de la martingala. Además, se dará una breve introducción a la teoría de medidas aleatorias. El contenido de este capítulo esta basado en [18, 19, 20, 8, 13, 14, 1, 17, 21, 10, 7].

1.1. Procesos de ramificación con espacio de estados continuo

Una manera de obtener un proceso de ramificación con espacio de estados continuo fue llevado a cabo por John Lamperti (1967). Para ello, se inicia recordando los proceos de ramificación más simples. El **proceso de Bienaymé-Galton-Watson** describe la evolución a tiempo discreto de una población de individuos quienes se reproducen asexualmente y de manera independiente al resto de la población de acuerdo a una misma distribución μ (o distribución de descendencia). Más precisamente, dado una distribución μ definido en los enteros no negativos, el proceso Bienaymé-Galton-Watson es la cadana de Markov $\{\mathbf{Z}_k, k \geq 0\}$ con valores en $\mathbb{N} \cup \{0\}$ tal que, condicionado a $\mathbf{Z}_n$, la siguiente igualdad en distribución se satisface

$$\mathbf{Z}_{n+1} = \sum_{i=1}^{\mathbf{Z}_n} \xi_i^{(n)}, \quad n = 1, 2 \ldots,$$

donde las variables $\{\xi_i^{(n)}; n, i \geq 1\}$ son independientes e identicamente distribuidas con distribución μ. Ahora, para cada $k \geq 1$, sea $\{\mathbf{Z}_n^k : n \geq 0\}$ una familia de procesos Bienaymé-Galton-Watson con población inicial z_k y distribución de descendencia π_k, la cual probablemente depende de k. Si existe una sucesión de constantes $a_k \uparrow \infty$ tal que el proceso

$$\left\{ \frac{1}{a_k} \mathbf{Z}_{[kt]}^k : t \geq 0 \right\} \tag{1.1.1}$$

converge, cuando $k \to \infty$, a un proceso estocástico límite $X = \{X_t, t \geq 0\}$, al menos en el sentido de convergencia débil de distribuciones finito dimensionales, entonces el proceso X

resulta ser un proceso de ramificación con espacio de estados continuo. Recíprocamente, cualquier proceso de ramificación con espacio de estados continuo puede ser obtenido de este modo.
Quizá el ejemplo más importante se da cuando en (1.1.1) se considera $\pi_k = \pi$ para toda k con π crítico ($\sum k\pi(k)$=1) y de varianza finita. En este caso, la convergencia se sostiene tomando $a_k = k$ y considerando la hipótesis adicional el limite de z_k/k, cuando $k \to \infty$, exista y sea positiva (ver [19]). Cuando la población es muy numerosa, en lugar de considerar el número de individuos conviene estudiar la "densidad"de individuos.
Como en el caso del proceso Bienaymé-Galton-Watson([18]), existe una manera de caracterizar a un porceso de ramificación con espacio de estados continuo (CSBP, por sus siglas en inglés) por medio de la transfomada de Laplace de las probabilidades de transición. Para llevar a cabo lo anterior, es necesario mencionar que el caso que interesa es el caso crítico, es decir, cuando el número esperado de descendientes de cada individuo es x, lo cual sucede si se cumple

$$\int p_t(x, dy)y = x.$$

Teorema 1.1.1. *([19])Se asume que la familia $(p_t(x, dy), t > 0, x \in \mathbb{R}_+)$ satisface las siguientes propiedades*

(i) $p_t(x, \cdot) * p_t(x', \cdot) = p_t(x + x', \cdot)$ *para todo* $t > 0, x, x' \in \mathbb{R}_+$.

(ii) $\int p_t(x, dy) \leq x$ *para toda* $t > 0, x \in \mathbb{R}_+$.

Si excluimos el caso trivial $p_t(x, \cdot) = \delta_0$ para toda $t > 0$ y $x \in \mathbb{R}_+$, entonces la transformada de Laplace de $p_t(x, dy)$ debe ser de la forma

$$\int p_t(x, dy)e^{-zy} = e^{-xu_t(z)}, \quad z \geq 0,$$

y la función $(u_t(z), t \geq 0, z \geq 0)$ es la solución no negativa de la ecuación integral

$$u_t(z) + \int_0^t \psi(u_s(z))ds = z,$$

con ψ una función de la forma

$$\psi(z) = \alpha z + \beta z^2 + \int \pi(dr)(e^{-rz} - 1 + rz),$$

donde $\alpha, \beta \geq 0$ y π es una medida $\sigma-$finita en $(0, \infty)$ tal que

$$\int \pi(dr)(r \wedge r^2) < \infty.$$

Definición 1.1.2. *El* ***$\psi-$proceso de ramificación con espacio de estados continuo*** *es el proceso de Markov en $\mathbb{R}_+$ $X = \{X_t, t \geq 0\}$ cuyas probabilidades de transición $P_t(x, dy)$ estan asociadas a la función ψ dada en el Teorema 1.1.1. La función ψ se llama el* ***mecanismo de ramificación*** *de X.*

Un caso especial es cuando $\psi(z) = \beta z^2$(el mecanismo de ramificación cuadrático). La ecuación integral presentada en el Teorema 1.1.1 resulta ser

$$u_t(z) + \int_0^t \beta(u_s(z))^2 ds = z.$$

el cual implica que $u_0(z) = z$. Observemos que, gracias al teorema fundamental del cálculo, esta ecuación se puede escribir como

$$\frac{\partial u_t(z)}{\partial t} + \beta u_t^2(z) = 0.$$

O equivalentemente,

$$\frac{\partial u_t(z)}{\partial t}\frac{1}{u_t^2(z)} + \beta = -\frac{\partial}{\partial t}\left(\frac{1}{u_t^{-1}(z)}\right) + \beta = 0.$$

Resolviendo esta última ecuación diferencial con condición inicial $u_0(z) = z$, obtenemos

$$\frac{1}{u_t^{-1}(z)} = \beta t + z,$$

es decir,

$$u_t(z) = \frac{z}{1 + \beta z t}.$$

Definición 1.1.3. *Al $\psi-$proceso de ramificación con espacio de estados continuo descrito en las lineas anteriores se llama* ***difusión de Feller*** *de parámetro β.*

Ejemplo 1.1.4. *Sea $\overline{\mathbf{N}}$ el proceso definido por la transformada de Laplace*

$$\mathbf{E}[\exp\{-\theta\overline{\boldsymbol{N}}_t\}|\overline{\boldsymbol{N}}_0 = 1] = \exp\left\{-\frac{\theta}{(1+\lambda t\theta)}\right\}, \quad \theta \geq 0.$$

Entonces, por el Teorema 1.1.1, $\overline{\mathbf{N}}$ es la difusión de Feller crítico de parámetro λ. Este proceso tiene deriva cero, varianza $2\lambda x$ y generador $L = \lambda x d^2/dx^2$ (ver, por ejemplo, [19, 1, 21]). Es bien sabido que $\overline{\mathbf{N}}$ puede ser visto como una aproximación de difusión de un proceso de nacimiento y muerte, crítico, lineal, sobre los enteros no negativos como se muestra a continuación.

Otra manera de obtener la difusión de Feller es por medio de ecuaciones diferenciales estocásticos (ver [19, 20]), esto es, dentando por B un movimiento Browniano unidemensional estándar, la difusión de Feller X se puede construir como la solución de la siguiente ecuación diferencial estocástica

$$dX_t = \sqrt{2\beta X_t} dB_t, \quad X_0 = x.$$

1.2. El Problema de la martingala

Una forma de caracterizar a los procesos de Markov fue llevado a finales de 1960 en [24]. El problema de la martingala fue usado como una forma de construir y estudiar procesos de Markov. Este método es particularmente útil cuando se trata de cuestiones de aproximación y convergencia en ley para procesos estocásticos, y destaca por proporicionar una herramienta para determinar si un proceso estocástico dado, bajo ciertas condiciones, satisface la propiedad de Markov. Una buena fuente de consulta sobre la teoría del problema de la martingala se encuentra en [8], que es donde se basa el resto de esta sección.
En adelante, L denota un espacio de Banach con norma $\|\cdot\|$. Se recuerda que un operador lineal acotado en L es una aplicación lineal $T : L \to L$ tal que $\|Tf\| \leq \|f\|$ para toda $f \in L$. Dado un operador lineal acotado B en L, se define, para cada $t \geq 0$,

$$\exp\{tB\} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k B^k.$$

Definición 1.2.1. *Una familia $\{T_t, t \geq 0\}$ de operadores lineales acotados en L se llama* ***semigrupo*** *si $T_0 = I$ y $T_{s+t} = T_s T_t$ para toda $s, t \geq 0$. Un semigrupo $\{T_t, t \geq 0\}$ se dice* ***fuertemente continuo*** *si $\lim_{t \to 0} T_t f = 0$ (en la norma $\|\cdot\|$) para toda $f \in L$; y es un semigrupo de* ***contracción*** *si $\|T_t\| \leq 1$ para toda $t \geq 0$.*

Proposición 1.2.2. *Sea $\{T_t, t \geq 0\}$ un semigrupo fuertemente continuo en L. Entonces, existe una constante $M \geq 1$ y $w \geq 0$ tal que*

$$\|T_t\| \leq M \exp\{wt\}.$$

Corolario 1.2.3. *Sea $\{T_t, t \geq 0\}$ un semigrupo fuertemente continuo en L. Entonces, para cada $f \in L$, la aplicación $t \mapsto T_t f$ es una función continua de $[0, \infty)$ a L.*

Se observa que si $S_t = T_t \exp\{-wt\}$ para cada $t \geq 0$, entonces $\{S_t, t \geq 0\}$ es un semigrupo fuertemente continuo en L tal que

$$\|S_t\| \leq M, \quad t \geq 0. \tag{1.2.1}$$

En particular, si $M = 1$, el semigrupo $\{S_t, t \geq 0\}$ es un semigrupo fuertemente continuo de contracción. También se observa que, si se define la norma $|||\cdot|||$ en L por

$$|||f||| = \sup_{t \geq 0} \|S_t f\|,$$

entonces $\|f\| \leq |||f||| \leq M \|f\|$ para cada $f \in L$, por lo que la nueva norma es equivalente a la norma original. En consecuencia, respecto a la norma $|||\cdot|||$, $\{S_t, t \geq 0\}$ es un semigrupo de contracción fuertemente continuo en L, lo que justifica que la mayoría de los resultados subsecuentes están en términos de semigrupos fuertemente continuos.
Ahora, dado un operador lineal(posiblemente no acotado) A en L, se denota por $\mathcal{D}(A)$ y $\mathcal{R}(A)$ el dominio y rango de A, respectivamente. Se define la **gráfica** de A por

$$\mathcal{G}(A) = \big\{(f, Af) \bigm| f \in \mathcal{D}(A)\big\} \subset L \times L.$$

Se nota que $L \times L$ es un espacio de Banach con suma componente a componente, multiplicación por escalar y norma

$$\|(f,g)\| = \|f\| + \|g\|.$$

Se dice que un operador lineal A es **cerrado** si $\mathcal{G}(A)$ es un subespacio cerrada de $L \times L$ y se dice **cerrable** si éste tiene una extensión lineal cerrado. Si A es cerrable, entonces la cerradura $\overline{A}$ de A es la mínima extensión lineal cerrada de A; mas específicamente, éste es el operador lineal cerrado B cuya gráfica es la cerradura (en $L \times L$) de la gráfica de A.

Definición 1.2.4. *El **generador infinitesimal** de un semigrupo $\{T_t, t \geq 0\}$ en L es el operador lineal A definido por*

$$Af = \lim_{t \to 0} \frac{1}{t}\Big(T_t f - f\Big). \tag{1.2.2}$$

El dominio de A, $\mathcal{D}(A)$, es el subespacio de todas las $f \in L$ para las cuales el límite (1.2.2) existe en norma $\|\cdot\|$.

Proposición 1.2.5. *Si A es el generador infinitesimal de un semigrupo fuertemente continuo $\{T_t, t \geq 0\}$ en L, entonces $\mathcal{D}(A)$ es denso en L y A es cerrado.*

Definición 1.2.6. *Un operador A en L se llama **disipativo** si*

$$\|\lambda f - Af\| \geq \lambda \|f\|$$

para toda $f \in \mathcal{D}(A)$ y $\lambda > 0$.

Teorema 1.2.7. *(**de Hille-Yoshida**) Un operador lineal A en L es el generador infinitesimal de un semigrupo de contracción fuertemente continuo en L si, y sólo si,*

(i) *$\mathcal{D}(A)$ es denso en L.*

(ii) *A es disipativo.*

(iii) *$\mathcal{R}(\lambda - \mathcal{A}) = L$ para algún $\lambda > 0$.*

Definición 1.2.8. *Se define el **generador total** $\hat{A}$ de un semigrupo de contracción $\{T_t, t \geq 0\}$ en por*

$$\hat{A} = \left\{ (f,g) \in L \times L \,\middle|\, T_t f - f = \int_0^t T_s g ds, \quad t \geq 0 \right\} \tag{1.2.3}$$

Proposición 1.2.9. *Sea L como antes y $\{T_t\}_{t\geq 0}$ un semigupo de contracción medible en L. Entonces el generador total $\hat{A}$ de $\{T_t\}_{t\geq 0}$ es lineal, disipativo y satisface*

$$(\lambda - \hat{A})^{-1} h = \int_0^\infty \exp\{-\lambda t\} T_t h dt$$

para toda $h \in \mathcal{R}(\lambda - \hat{A})$ y $\lambda > 0$. Si

$$T_s \int_0^\infty \exp\{-\lambda t\} T_t h dt = \int_0^\infty \exp\{-\lambda t\} T_{s+t} h dt$$

para toda $h \in L$, $\lambda > 0$ y $s \geq 0$, entonces $\mathcal{R}(\lambda - \hat{A}) = L$ para toda $\lambda > 0$.

Ahora, sea (E, r) un espacio métrico, $M(E)$ el espacio de todas las funciones real-valuadas y Borel medibles en E y $\mathcal{C}_b(E) \subset M(E)$ el espacio de Banach de funciones continuas y acotadas con la norma $\|f\| = \sup_{x\in E} |f(x)|$.
Una manera de caracterizar a un proceso de Markov es por medio de sus probabilidades de transición; no obstante, este método, en general, no es muy útil debido a la dificultad para obtener de manera explícita las probabilidades de transición. Esto motiva el siguiente resultado.

Proposición 1.2.10. *Sean E separable y X un proceso de Markov con valores en E y condición inicial ν correspondiente a un semigrupo $\{T_t\}_{t\geq 0}$ en un subespacio cerrado $L \in \mathcal{C}_b(E)$ con $\|f\| = \sup_{x\in E} |f(x)|$. Si L es separante[1], entonces $\{T_t\}_{t\geq 0}$ y ν determinan las distribuciones finito dimensionales de X.*

Como las distribuciones finito dimensionales de un proceso de Markov están determinadas por el semigrupo correspondiente $\{T_t\}_{t\geq 0}$, ellas están determinadas por su generador total $\hat{A}$ o por un conjunto suficientemente grande $A \subset \hat{A}$. Uno de los mejores enfoques para determinar cuando un conjunto es sufiecientemente grande es a través del **problema de la martingala**, el cual esta basado en la siguiente proposición.

Proposición 1.2.11. *Sea X un proceso de Markov progresivo (para cada $t \geq 0$, la restricción de X a $[0, t] \times \Omega$ es $\mathcal{B}([0, t]) \times \mathcal{F}_t-$medible) con probabilidades de transición $P(t, x, \Gamma)$ y sea $\{T_t\}_{t\geq 0}$ y $\hat{A}$ como antes. Si $(f, g) \in \hat{A}$ entonces*

$$f(X_t) - \int_0^t g(X_s)ds \tag{1.2.4}$$

es una $\mathcal{F}_t^X$-martingala.

Lo que hace realmente importante al problema de la martingala, es el recíproco del resultado anterior. Básicamente, este afirma que si un proceso, bajo ciertas condidiciones, satisface que la expresión (1.2.4) es una martingala, entonces es un proceso de Markov. Como se mencionó, se usará las ideas desarrolladas por Stroock y Varadhan usando la propiedad de martingala como medio de caracterizar los procesos de Markov asociados con un generador A dado. Sea (E, r) un espacio métrico. Ocacionalmente, A denotará un operador multivariado, por lo que se puede pensar a A como un subconjunto de $\mathcal{C}_b(E) \times \mathcal{C}_b(E)$(no necesariamente lineal).

Definición 1.2.12. *Una **solución del problema de la martingala para** A es un proceso estocástico medible X con valores en E definido en algún espacio de probabilidad $(\Omega, \mathcal{F}, \mathbf{P})$ tal que para $(f, g) \in A$, (1.2.4) es una martingala con respecto a la filtración*

$${}^*\mathcal{F}_t^X := \mathcal{F}_t^X \bigvee \sigma \left\langle \int_0^s h(X_u)du \mid s \leq t, h \in \mathcal{C}_b(E) \right\rangle,$$

donde $\mathcal{C}_b(E)$ es el espacio de Banach de funciones acotadas con $\|f\| = \sup_{x\in E} |f(x)|$.

---

[1]En el sentido usual, a saber, M es una clase separante siempre que $\mathbf{P}, \mathbf{Q} \in \mathcal{P}(E)$ y $\displaystyle\int f d\mathbf{P} = \int f d\mathbf{Q}$, $f \in M$, se cumple $\mathbf{P} = \mathbf{Q}$.

Si $\{\mathcal{G}_t\}_{t\geq 0}$ es una filtración con $\mathcal{G}_t \supset^* \mathcal{F}_t^X$ para toda $t \geq 0$, y (1.2.4) es una $\mathcal{G}_t-$martingala para toda $(f,g) \in A$, se dice que X es una solución al problema de la martingala para A con respecto a $\{\mathcal{G}_t\}_{t\geq 0}$. Cuando se especifica una distribución inicial $\mu \in \mathcal{P}(E)$, se dice que una solución X al problema de la martingala para A es una solución al problema de la martingala para (A,μ) si la $\mathbf{P}X_0^{-1} = \mu$.
Usualmente, X tiene trayectorias en $\mathcal{D}_E[0,\infty)$. Con esto en mente, es conveniente llamar a un medida de probabilidad $\mathbf{P} \in \mathcal{P}(\mathcal{D}_E[0,\infty))$ una solución al problema de la martingala para A(o para (A,μ)) si el proceso canónico definido en $(\mathcal{D}_E[0,\infty), \mathcal{F}_E, \mathbf{P})$ por

$$X_t(w) := \omega(t), \quad \omega \in \mathcal{D}_E[0,\infty), \quad t \geq 0,$$

es una solución al problema de la martingala para A(o para (A,μ)) como se definió anteriormente.
Se dice que la solución al problema de la martingala para (A,μ) es **única** si cualesquiera dos soluciones tienen las mismas ditribuciones finito dimensionales. Si existe una única solución al problema de martingala para (A,μ), se dice que el problema de la martingala para (A,μ) está **bien planteado.** Si esto es válido para toda $\mu \in \mathcal{P}(E)$ entonces se dice que el problema de la martingala para A está bien planteado. Finalmente, se dice que el problema de la martingala para (A,μ) está bien de planteado en $\mathcal{D}_E[0,\infty)$ si existe una única solución $\mathbf{P} \in \mathcal{P}(E)$. El siguiente teorema establece que un proceso de Markov es la única solución del problema de la martingala para su generador infinitesimal.

Teorema 1.2.13. *Sea E separable, y $A \in \mathcal{C}_b(E) \times \mathcal{C}_b(E)$ lineal y disipativo. Se asume, además, que existe $A' \subset A$ tal que A' es lineal y satisface $\overline{\mathcal{R}(\lambda - A')} = \overline{\mathcal{D}(A')} \equiv L$ con L separable y algún $\lambda > 0$. Sea $\mu \in \mathcal{P}(E)$ y supóngase que X es una solución al problema de martingala de (A,μ). Entonces, X es un proceso de Markov correspondiente al semigrupo en L generado por la cerradura de A', y la unicidad se sostiene para el problema de martingala para (A,μ).*

1.3. Medidas aleatorias

Sea (E,τ) un espacio topológico. Se dice que (E,τ) es **metrizable** si existe una métrica d en E tal que τ es inducida por las bolas abiertas

$$B_\epsilon(x) := \{y \in E \mid d(x,y) < \epsilon\}.$$

En adelante, E denota un espacio topológico Hasudorff, localmente compacto y segundo numerable. A tal espacio se le conoce como **espacio Polaco**, es decir, existe una métrica en E que lo hace un espacio completo y separable y que hace metrizable a la topología en E. Se denota por $\mathcal{E}$ a la $\sigma-$álgebra de Borel E, es decir, la $\sigma-$álgebra generada por los conjuntos abiertos de E. Se recuerda que un subconjunto de $\mathcal{E}$ es **acotado** (o **relativamente compacto**), si su cerradura es compacta; sea $\widehat{\mathcal{E}}$ la familia de tales conjuntos.

Definición 1.3.1. *Una medida μ definida en el espacio $(E,\mathcal{E})$ se dice que es una **medida Borel-Radon**(o localmente finita) si $\mu(B) < \infty$, para todo $B \in \widehat{\mathcal{E}}$.*

Sea $\mathcal{M}(E)$ al espacio de medidas en $(E,\mathcal{E})$ que son Borel-Radon. Para cualquier $\mu \in \mathcal{M}(E)$ se define $\hat{\mathcal{E}}_\mu = \{B \in \hat{\mathcal{E}} \mid \mu(\partial B) = 0\}$. El espacio $\mathcal{M}(E)$ contiene subespacios con propiedades interesantes.

$$\begin{aligned}
\mathcal{M}_f(E) &:= \left\{\mu \in \mathcal{M}(E) \,\middle|\, \mu \text{ es finita } \right\},\\
\mathcal{M}_1(E) &:= \left\{\mu \in \mathcal{M}(E) \,\middle|\, \mu \text{ es medida de probabilidad } \right\},\\
\mathcal{N}(E) &:= \left\{\mu \in \mathcal{M}(E) \,\middle|\, \mu \text{ toma valores en } \mathbf{Z}^+ \right\},\\
\mathcal{M}_T(E) &:= \left\{\mu \in \mathcal{M}(E) \,\middle|\, \text{ existe } p \geq 0 \text{ tal que } \int (1+|x|^2)^{-p}\mu(dx) < \infty \right\}.
\end{aligned}$$

Al espacio $\mathcal{M}_T(E)$ se le conoce como el **espacio de medidas temperadas** y fue introducido por Iscoe en [10] para $E = \mathbb{R}^d$, con el propósito de estudiar resultados sobre procesos con valores en espacios de medidas de Radon no finitas, como por ejemplo la medida de Lebesgue en $\mathbb{R}^d$. Es un hecho bien conocido que la medida de Lebesgue en $\mathbb{R}^d$ pertenece a $\mathcal{M}_p(\mathbb{R}^d)$ para $p > \frac{d}{2}$ (ver [10]). Sean $\mathcal{C}_b(E)$ la familia de funciones $f : E \to \mathbb{R}$ con continuos y acotadas, y $\mathcal{C}_K(E) \subset \mathcal{C}_b(E)$ el subespacio de funciones con soporte compacto.

Definición 1.3.2. *(i) Sea $\mu, \mu_1, \ldots \in \mathcal{M}_f(E)$. Se dice que $\{\mu_n\}_{n\geq 1}$ **converge débilmente** a μ, denotado por $\mu_n \Rightarrow \mu$, si*

$$\int f d\mu_n \to \int f d\mu, \quad n \to \infty,$$

para toda $f \in \mathcal{C}_b(E)$.
***(ii)** Sea $\mu, \mu_1, \ldots \in \mathcal{M}(E)$. Se dice que $\{\mu_n\}_{n\geq 1}$ **converge vagamente** a μ, denotado por $\mu_n \xrightarrow{v} \mu$, si*

$$\int f d\mu_n \to \int f d\mu, \quad n \to \infty,$$

para toda $f \in \mathcal{C}_K(E)$.

Se escribe π_B y π_f para denotar los mapeos $\mu \mapsto \mu B$ y

$$\mu \mapsto \mu f = \langle \mu, f\rangle := \int f d\mu$$

respectivamente. En análisis funcional la convergencia débil es llamada convergencia débil $*$. La convergencia débil induce en $\mathcal{M}_f(E)$ la **topología débil**, esto es, la topología generada por los mapeos π_f, $f \in \mathcal{C}_b(E)$, y se denota por τ_w. Si E es separable, $(\mathcal{M}_f(E), \tau_w)$ es metrizable. Similarmente, la **topología vaga** en $\mathcal{M}(E)$ está generada por los mapeos π_f, $f \in \mathcal{C}_c(E)$, y se denota por τ_ν. Se denotará por $\mathfrak{M}(E)$ a la $\sigma-$**álgebra generada por la topología vaga**. Si E es localmente compacto, entonces $(\mathcal{M}(E), \tau_\nu)$ es Hausdorrf, y si E es un espacio Polaco, entonces $(\mathcal{M}(E), \tau_\nu)$ es nuevamente un espacio Polaco [13].
Se puede definir una métrica en $\mathcal{M}(E)$ de la siguiente manera. Sean $f_1, f_2, \ldots$ denso en $\mathcal{C}_K^+(E)$

$$\rho(\mu,\nu) = \sum_{k\geq 0} 2^{-k} \left(|\mu f_k - \nu f_k| \wedge 1\right), \quad \mu,\nu \in \mathcal{E}.$$

Entre las propiedades fundamentales de la topología vaga resaltan las siguientes: una sub-base topológica para $\mathcal{M}(E)$ está formada por todas las interseccione finitas de los conjuntos

$$\left\{\mu \,\middle|\, a < \mu f < b\right\}, \; 0 < a < b, \; f \in \mathcal{C}_c(E);$$

una vecindad básica de $\mu \in \mathcal{M}(E)$ bajo esta topología son conjuntos de la forma

$$\left\{\nu \,\middle|\, s_i < \mu f_i - \nu f_i < t_i\right\}, \; f_i \in \mathcal{C}_c(E), \; s_i, t_i \in \mathbb{R}, \; i = 1, \ldots, m, \; m = 1, 2, \ldots.$$

Teorema 1.3.3. *Para cualquier espacio E localmente compacto, segundo numerable Hausdorff se tiene*

(i) $(\mathcal{M}(E), \rho)$ *es un espacio Polaco respecto a la topología vaga;*

(ii) *un conjunto $A \subset \mathcal{E}$ es relativamente compacto respecto a la topología vaga si, y sólo si,* $\sup_{\mu \in A} \mu f < \infty$ *para toda $f \in \mathcal{C}_K^+$;*

(iii) *si $\mu_n \xrightarrow{v} \mu$ y $B \in \hat{\mathcal{E}}$ con $\mu \partial B = 0$, entonces $\mu_n B \to \mu$;*

(iv) *$\mathcal{B}(\mathcal{M}(E))$ está generada por los mapeos π_f, $f \in \mathcal{C}_K^+$, y también para cualquier $m \in \mathcal{M}(E)$ por los mapeos π_B con $\hat{\mathcal{E}}_m$.*

Definición 1.3.4. *Sea $(\Omega, \mathcal{F}, \mathbf{P})$ un espacio de probabilidad. Una **medida aleatoria** ξ en E es una función medible*

$$\xi : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathcal{M}(E), \mathfrak{M}(E)).$$

*La **distribución** o **ley** de ξ, denotado por $\mathbf{P}\xi^{-1}$, es la medida definida en $(\mathcal{M}(E), \mathfrak{M}(E))$ dada por*

$$\mathbf{P}\xi^{-1}(B) = \mathbf{P}\left(\xi^{-1}(B)\right) = \mathbf{P}(\xi \in B), \; \forall B \in \mathfrak{M}(E).$$

Se observa que dada la medida aleatoria ξ, para cada $\omega \in \Omega$, se obtiene una medida localmente finita en E. Tal medida será denotada por $\xi(\omega, \cdot)$. Esperando no causar confusión, se escribirá $\xi(\cdot)$ en vez de $\xi(\omega, \cdot)$, así por ejemplo, la medida $\xi(\omega, \cdot)$ en $A \in \mathcal{E}$ se escribirá $\xi(A)$ en vez de $\xi(\omega, A)$. También, se observa que la medida $\mathbf{P}\xi^{-1}$ se encuentra definida en un espacio topológico $\mathcal{M}$.

Definición 1.3.5. *La **intensidad** de la medida aleatoria ξ es la medida*

$$(\mathbf{E}\xi)(B) = \mathbf{E}[\xi(B)] = \int_\Omega \xi(\xi, B) d\mathbf{P}, \quad B \in \mathfrak{M}(E).$$

Como es usual, dado un elemento aleatorio, es conveniente trabajar con su transformada de Laplace o su función característica. Dada una medida aleatoria, se define el **funcional de Laplace** y el **funcional característico** de ξ por

$$\mathbf{E}[\exp\{-\langle \psi, \xi \rangle\}], \; \psi \in \mathcal{C}_c(E)$$

y

$$\mathbf{E}[\exp\{i \langle \psi, \xi \rangle\}], \; \psi \in \mathcal{C}_c(E),$$

respectivamente. Aquí $\mathbf{E}$ denota el valor esperado con respecto a la ley de ξ. Como es usual, cada uno de los funcionales anteriores determina de manera única la ley de ξ.

Teorema 1.3.6. *Las medidas de probabilidad en $(\mathcal{M}(E), \mathfrak{M}(E))$ quedan determinadas, unívocamente, por su funcional de Laplace o por su funcional característico.*

Ejemplo 1.3.7. *Una medida aleatoria N, es una* ***medida aleatoria de Poisson*** *si*

(i) *para $F \in \xi$*

$$\mathbf{P}\left(N(F)=k\right)=\exp\big\{-\mu(F)\big\}\frac{\mu(F)^k}{k!}, \quad k \geq 0,$$

donde μ es la medida de intensidad de N;

(ii) *para toda $k \geq 1$, $F_1, \ldots, F_k \in \xi$ con $F_i \cap F_j = \emptyset$, $i \neq j$, entonces*

$$N(F_1), \ldots, N(F_k)$$

son independientes.

Por el Teorema 1.3.6, el funcional de Laplace de la medida aleatoria de Poisson está dado por

$$\mathbf{E}\left[\exp\{-\langle \psi, P\rangle\}\right]=\exp\left\{\int\big(\exp\{-\psi(x)\}-1\big)\mu(dx)\right\}, \quad \psi \in \mathcal{C}_K^+.$$

Definición 1.3.8. *Sea $(\Omega, \mathcal{F}, \mathbf{P})$ un espacio de probabilidad. Un* ***proceso estocástico con valores en*** *$\mathcal{M}(E)$, o proceso estocástico con valores en medidas, es una familia $X = \{X_t\}_{t\geq 0}$, es decir, para cada $t \geq 0$,*

$$X_t : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathcal{M}(E), \mathfrak{M}(E))$$

es una medida aleatoria. El valor esperado condicionado a $X_0 = \mu$ se denota por $\mathbf{E}_\mu$. El proceso estocástico se dice proceso de Markov si X es un proceso de Markov con espacio de estados $(\mathcal{M}(E), \mathfrak{M}(E))$.

Sea X un proceso estocástico definido en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbf{P})$ y con valores en $(\mathcal{M}(E), \mathfrak{M}(E))$, se define su **familia de distribuciones finito dimensionales** a la familia de medias de probabilidad

$$\{\mu_{t_1,\ldots,t_n} \,\big|\, 0 \leq t_1 \leq \ldots \leq t_n, n \geq 1\},$$

donde, para cada $n \geq 1$ y $t_1 \leq \ldots \leq t_n$, la medida $\mu_{t_1,\ldots,t_n}$ es una medida de probabilidad en $\mathfrak{M}(E)^n$ dada por

$$\mu_{t_1,\ldots,t_n}(\Gamma)=\mathbf{P}\left((X_{t_1},\ldots,X_{t_n})\in\Gamma\right), \ \Gamma \in \mathfrak{M}(E)^n.$$

Un **proceso de ramificación** con valores en medidas es un proceso de Markov X con espacio de estados $\mathcal{M}(E)$ y probabilidades de transición $\{P_t\}_{t\geq 0}$ que satisface la **propiedad de ramificación**

$$P_t(\cdot, \mu_1+\mu_2)=P_t(\cdot,\mu_1) * P_t(\cdot,\mu_2), \ \mu_1, \mu_2 \in \mathfrak{M}(E), t \geq 0,$$

donde $*$ denota la convolución de medidas.

Capítulo 2

Convergencia de procesos estocásticos

En este capítulo se estudia la convergencia de procesos estocásticos, tomando como punto de partida la teoría de convergencia débil de medidas de probabilidad en espacio métricos, la cual es analizada en la primera sección. Entre los resultados establecidos resalta la caracterización de los subconjuntos compactos. Posteriormente, en la segunda sección se hace un breve estudio del espacio de funciones real valuadas en $[0,1]$ continuas por la derecha y con límites por la izquierda. Este estudio se extenderá al considerar funciones con valores en $\mathbb{R}^+$ y valores un espacio métrico en la tercera sección. Finalmente, en la última sección, se da un criterio de tensión para procesos estocásticos con trayectorias en un espacio de medidas. El contenido de este capítulo está basado en [8, 4, 22, 25, 11].

2.1. Medidas en espacios métricos

En adelante, (E, r) denota un espacio métrico. Sea $\mathcal{E}$ la $\sigma-$álgebra de Borel generado por los conjuntos abiertos (o cerrados) de E. Una medida de probabilidad en $\mathcal{E}$ es una función de conjuntos no negativa, numerablemente aditiva y satisface que $\mathbf{P}(E) = 1$.

2.1.1. Convergencia débil de medidas

Teorema 2.1.1. *Toda medida de probabilidad* $\mathbf{P}$ *en* $(E, \mathcal{E})$ *es* ***regular****, es decir, para toda* $A \in \mathcal{E}$ *y todo* $\epsilon > 0$ *existe un conjunto cerrado* F *y un conjunto abierto* G *tal que* $F \subset A \subset G$ *y* $\mathbf{P}(G - F) < \epsilon$.

Demostración. Sea $\mathcal{G}$ la familia de conjuntos de $\mathcal{E}$ que cumplen con la propiedad deseada. Se quiere probar que $\mathcal{G} = \mathcal{E}$. Con esto en mente, es suficiente demostrar que $\mathcal{G}$ es una $\sigma-$álgebra y que contiene a la familia de conjuntos cerrados de S. Primero, se probará el primer caso.

- Como $\emptyset$ y E son conjuntos abiertos y cerrados, $\emptyset$ y E pertenecen a $\mathcal{G}$.

- Sea $A \in \mathcal{G}$. Entonces, dado $\epsilon > 0$, existe un conjunto cerrado F_ϵ y un conjunto abierto G_ϵ tales que $F_\epsilon \subset A \subset G_\epsilon$ y $\mathbf{P}(G_\epsilon - F_\epsilon) < \epsilon$. Luego, $G'_\epsilon \subset A' \subset F'_\epsilon$ y $\mathbf{P}(F'_\epsilon - G'_\epsilon) < \epsilon$ con G'_ϵ cerrado y F'_ϵ abierto. Esto es, $A' \in \mathcal{G}$, por lo que $\mathcal{G}$ es cerrado bajo complementos.

- Sea $\{A_n\}_{n\geq 1} \subset \mathcal{G}$ y $A = \bigcup_{n\geq 1} A_n$. Además, sea $\epsilon > 0$. Entonces, para cada $n \in \mathbb{N}$, existe un conjunto cerrado $F_{n,\epsilon}$ y un conjunto abierto $G_{n,\epsilon}$ tales que $F_{n,\epsilon} \subset A_n \subset G_{n,\epsilon}$ y $\mathbf{P}\left(F_{n,\epsilon} - G_{n,\epsilon}\right) < \epsilon/2^{n+1}$. Sea $G_\epsilon = \bigcup_{n\geq 1} G_{n,\epsilon}$ y $F = \bigcup_{n\geq 1} F_{n,\epsilon}$. Ya que $\mathbf{P}$ es una medida, se puede elegir $k \in \mathbb{N}$ tal que

$$\mathbf{P}\left(C - \bigcup_{n=1}^{k} F_{n,\epsilon}\right) < \epsilon.$$

Sea $F_\epsilon = \bigcup_{n=1}^{k} F_{n,\epsilon}$. Entonces, G_ϵ es abierto, F_ϵ es cerrado, $F_\epsilon \subset A \subset G_\epsilon$ y

$$\begin{aligned}
\mathbf{P}\left(G_\epsilon - F_\epsilon\right) &\leq \mathbf{P}\left(G_\epsilon - C\right) + \mathbf{P}\left(C - F_\epsilon\right) \\
&\leq \sum_{n\geq 1} \mathbf{P}\left(U_{n,\epsilon} - F_{n,\epsilon}\right) + \epsilon/2 \\
&= \sum_{n\geq 1} \epsilon/2^{n+1} + \epsilon/2 \\
&= \epsilon.
\end{aligned}$$

Por lo tanto, $A \in \mathcal{G}$, es decir, $\mathcal{G}$ es cerrado bajo unión numerable. Esto finaliza la prueba de que $\mathcal{G}$ es una $\sigma-$álgebra.

Ahora, para $A \subset E$ y $\epsilon > 0$ se definen

$$r(x, A) := \text{ínf}\{r(x, y) \,\big|\, y \in A\}, \tag{2.1.1}$$

y

$$A^\delta := \{x \in S : r(x, A) < \delta\}. \tag{2.1.2}$$

Así, si A es un conjunto cerrado de E, entonces, eligiendo $F = A$ y $G = A^\delta$ para algún $\delta > 0$[1] y usando la continuidad de la medida, se cumple $F \subset A \subset G$ tal que $\mathbf{P}(G - F) < \epsilon$. Por lo que la familia de conjuntos cerrados en E está contenida $\mathcal{G}$. Así, por lo probado anteriormente, $\mathcal{E} = \mathcal{G}$. ■

La importancia de este teorema radica en que la medida $\mathbf{P}$ está completamente determinada por los valores de $\mathbf{P}(F)$ con F conjunto cerrado en el siguiente sentido: si $\mathbf{P}$ y $\mathbf{Q}$ son medidas de probabilidad en E tales que $\mathbf{P}(F) = \mathbf{Q}(F)$ para todo conjunto cerrado F, entonces $\mathbf{P}(A) = \mathbf{Q}(A)$ para todo $A \in \mathcal{E}$. Esto es, la familia de conjuntos cerrados de E forman una clase separante.

Definición 2.1.2. *Sea $\mathcal{P}(E)$ la familia de medidas de probabilidad en E. Una colección de conjuntos $\mathcal{A} \subset \mathcal{E}$ se llama **clase separante** si para $\mathbf{P}, \mathbf{Q} \in \mathcal{P}(E)$ tales que $\mathbf{P}(A) = \mathbf{Q}(A)$, para toda $A \in \mathcal{A}$, implica que $\mathbf{P} = \mathbf{Q}$.*

Se llama una clase separante ya que los valores de $\mathbf{P}$ en la clase $\mathcal{A}$ son suficientes para separar $\mathbf{P}$ de otras medidas de probabilidad en S. Se recuerda que una familia de conjuntos de $\mathcal{E}$, A, es una clase separante si éste es un $\pi-$sistema y genera [2] a $\mathcal{E}$.

[1] La elección de un δ es posible ya que A^δ decrece a A cuando δ decrece a 0.

[2] La prueba se sigue del Lema de Dynkin $\pi - \lambda$, a saber, si A es un $\pi-$sistema tal que $A \subset \mathcal{L}$, con $\mathcal{L}$ un $\lambda-$sistema, entonces $\sigma(A) \subset \mathcal{L}$.

Ejemplo 2.1.3. *La clase de conjuntos cerrados forman una clase separante en todo espacio métrico.*

Ejemplo 2.1.4. *Sea* $\mathbb{Q}_p$ *el espacio de números* $p-$*ádicos con norma* $|\cdot|_p$ *(la norma* $p-$*ádica), donde* p *es un número primo. Por el ejemplo anterior, la clase de conjuntos cerrados, inducidos por la norma* $p-$*ádica, forman una clase separante. Además, se puede probar que* $\mathbb{Q}_p$ *es un espacio completo y separable.*

Ejemplo 2.1.5. *Sea* $\mathbb{R}^k$ *el espacio euclidiano* $k-$*dimensional con la metrica usual* $|\cdot|$, *y* $\mathcal{R}^k$ *la* $\sigma-$*álgebra de Borel* $k-$*dimensional(la generada por los abiertos en* $\mathbb{R}^k$*). La familia de conjuntos de la forma* $\{(y_1,\ldots,y_k) : y_i \leq x_i, i \leq k\}$ *forman un* $\pi-$*sistema que genera* $\mathcal{R}^k$, *y en consecuencia, constituyen una clase separante. Como en ejemplo anteriore, resulta que* $\left(\mathbb{R}^k, |\cdot|\right)$ *es un espacio completo y separable.*

Ejemplo 2.1.6. *Sea* $\mathbb{R}^\infty$ *el espacio de sucesiones* $x = (x_1, x_2, \ldots)$ *de número reales. Si* $b(y,z) = 1 \wedge |y-z|$, *entonces* b *es una métrica en* $\mathbb{R}$ *y equivalente a la norma usual, y bajo ésta,* $\mathbb{R}$ *es completo y separable. Sea* $\overline{d} : \mathbb{R}^\infty \to \mathbb{R}^+$ *definido por*

$$\overline{d}(x,y) = \sum_{i\geq 1} b(x_i,y_i)/2^i$$

y $\pi_k : \mathbb{R}^\infty \to \mathbb{R}$ *la* ***proyección canónica****, es decir, para* $x \in \mathbb{R}^\infty$,

$$\pi_k(x) = (x_1,\ldots,x_k).$$

Entonces, la familia $\mathcal{R}_f^\infty$ *dado por*

$$\mathcal{R}_f^\infty = \{\pi_k^{-1}(H) : H \in \mathcal{R}^k, k \geq 1\}.$$

Este conjunto resulta ser una clase separante pues es un $\pi-$*sistema y genera a* $\mathcal{R}^\infty$*. Se puede probar que* $\left(\mathbb{R}^\infty, \overline{d}\right)$ *es un espacio métrico completo y separable.*

Ejemplo 2.1.7. *Sea* $C = C[0,1]$ *el espacio de funciones continuas en* $[0,1]$*. Se define la* ***norma uniforme*** *en* C *por*

$$\|f\| = \sup_{t\in[0,1]}\{|x(t)|\}.$$

Se denota por ρ *a la métrica inducida por esta norma en* C, *y por* $\mathcal{C}$ *a la* $\sigma-$*álgebra de Borel de* C*. Ahora, para* $0 \leq t_1 \leq \cdots \leq t_k \leq 1$, *se define la proyección canónica de* C *a* $\mathbb{R}^k$ *como sigue*

$$\pi_{t_1,\ldots,t_k}(f) = (x(t_1),\ldots,x(t_k)).$$

Entonces, la clase de conjuntos de la forma

$$\mathcal{C}_f := \{\pi^{-1}_{t_1,\ldots,t_k}(H) : 0 \leq t_1 \leq \ldots \leq t_k \leq 1, H \in \mathcal{R}^k\}$$

es un $\pi-$*sistema que genera a* $\mathcal{C}$, *y en consecuencia, es una clase separante. Se puede probar que* (C,ρ) *es un espacio métrico completo y separable.*

Otra manera de determinar cuándo dos medidas coinciden es usando funciones continuas y acotadas. Esta relación motiva la definición de convergencia débil de medidas. Sea

$$\mathcal{C}_b(E) := \{f : E \to \mathbb{R} \mid f \text{ es continua y acotada}\}.$$

Se denota por $\mathcal{P}(E)$ el espacio de medidas de probabilidad definidas en E y se define

$$\langle \mathbf{P}, f\rangle := \int f d\mathbf{P}, \quad \mathbf{P} \in \mathcal{P}(E), \ f \in \mathcal{C}_b(E).$$

El siguiente resultado muestra que $\mathbf{P}$ también está determinado por los valores de $\langle \mathbf{P}, f\rangle$ para funciones f continuas y acotadas

Teorema 2.1.8. *Sean $\mathbf{P}$ y $\mathbf{Q}$ medidas de probabilidad en $\mathcal{E}$. Si $\langle \mathbf{P}, f\rangle = \langle \mathbf{Q}, f\rangle$ para toda función real valuada, uniformemente continua y acotada, entonces $\mathbf{P}$ y $\mathbf{Q}$ coinciden en $\mathcal{E}$.*

Demostración. Sea F un conjunto cerrado y $\epsilon > 0$. Se define $f : E \to [0, 1]$ por

$$f(x) = (1 - r(x, F)/\epsilon)^+.$$

donde $(x)^+ = x \vee 0$ y $r(x, F)$ definida como en (2.1.1). Se observa que f satisface

$$|f(x) - f(y)| \leq r(x, y)/\epsilon,$$

es decir, f es una función uniformemente continua (a modo de ejemplo, si $x \in F$ y $y \notin F$ con $r(y, F)/\epsilon \leq 1$, entonces $|f(x) - f(y)| = r(y, F)/\epsilon \leq r(y, x)/\epsilon$). También, se observa que $x \in F$ implica que $f(x) = 1$, mientras que $x \notin F^\epsilon$ implica que $r(x, F) \geq \epsilon$ (y en consecuencia, $r(x, F)/\epsilon \geq 1$) y así $f(x) = 0$. Así

$$\mathbf{1}_F(x) \leq f(x) \leq \mathbf{1}_{F^\epsilon}(x). \tag{2.1.3}$$

Ahora, sea F un conjunto cerrado. Entonces, por hipótesis y (2.1.3), se cumple que

$$\mathbf{P}(F) \leq \langle \mathbf{P}, f\rangle = \langle \mathbf{Q}, f\rangle \leq \mathbf{Q}F^\epsilon.$$

Luego, la continuidad de $\mathbf{Q}$ y la cerradura de F implican, cuando $\epsilon \downarrow 0$, $\mathbf{P}(F) \leq \mathbf{Q}(F)$. Por simetría y el Teorema 2.1.1 se tiene $\mathbf{P} = \mathbf{Q}$. ■

Definición 2.1.9. (i) *Sean $\mathbf{P}_n$ y $\mathbf{P}$ medidas de probabilidad en $(E, \mathcal{E})$. Se dice que la sucesión de medidas $\mathbf{P}_n$ **converge débilmente** a $\mathbf{P}$ en E, denotado por $\mathbf{P}_n \Rightarrow \mathbf{P}$, si*

$$\langle \mathbf{P}_n, f\rangle \to \langle \mathbf{P}, f\rangle,$$

cuando $n \to \infty$, para toda función real valuada f continua y acotada en E.
(ii) *Se dice que un conjunto A es un **conjunto de continuidad** $\mathbf{P}$, o de $\mathbf{P}-$continuidad, si $\mathbf{P}(\partial A) = 0$.*

Una consecuencia de los Teoremas 2.1.1 y 2.1.8 es la justificación de la posibilidad de trabajar con medidas $\mathbf{P}(A)$ o $\langle \mathbf{P}, f\rangle$, siendo el uso de esta última el mas simple o natural. Antes de continuar, conviene dar algunos ejemplos.

Ejemplo 2.1.10. *Sea $\delta_x(A)$ la delta de Dirac en $x \in E$. Si $x_n \to x_0$, entonces, para toda función continua f, se tiene*

$$\langle \delta_{x_n}, f \rangle = f(x_n) \to f(x_0) = \delta_{x_0}.$$

Por lo que $\delta_{x_n} \Rightarrow \delta_{x_0}$. El recíproco también es cierto (método contrapositiva). Por lo tanto, $\delta_{x_n} \Rightarrow \delta_{x_0}$ si, y sólo si, $x_n \to x_0$.

Ejemplo 2.1.11. *Sea $E = [0,1]$, $\mathbf{P}$ la medida de Lebesgue en E y $\mathbf{P}_n$ en E como*

$$\mathbf{P}_n = \frac{1}{m_n} \sum_{1 \leq k \leq m_n} \delta_{x_{n,k}},$$

donde los puntos $x_{n,k}$ son tales que, para cualquier $J \subset [0,1]$,

$$\mathbf{P}_n(J) = \frac{\sharp\{k : x_{n,k} \in J\}}{m_n} \to \mathbf{P}(J)$$

cuando $n \to \infty$. Aqui $\sharp$ denota la cardinalidad de un conjunto. Esta condición es suficiente para obtener $\mathbf{P}_n \Rightarrow \mathbf{P}$. Este hecho se puede probar usando sumas de Riemann. En efecto, sea $J_1, \ldots J_r$ una partición de E. Entonces, para f una función continua, se tiene

$$\langle \mathbf{P}_n, f \rangle = \frac{1}{m_n} \sum_{1 \leq k \leq m_n} f(x_{n,k}) \leq \frac{1}{m_n} \sum_{1 \leq i \leq r} \sup\{f(x) : x \in J_i\} \sharp\{k : x_{n,k} \in J_i\}.$$

Luego, por hipótesis y definición de la integral(por medio de sumas de Riemann),

$$\frac{1}{m_n} \sum_{1 \leq i \leq r} \sup\{f(x) : x \in J_i\} \sharp\{k : x_{n,k} \in J_i\} \to \sum_{1 \leq i \leq r} \sup\{f(x) : x \in J_i\} \mathbf{P}(J_i) \leq \langle \mathbf{P}, f \rangle,$$

cuando $n \to \infty$. Esto implica $\langle \mathbf{P}_n, f \rangle \leq \langle \mathbf{P}, f \rangle$. La otra desigualdad se obtiene consiederando las sumas inferiores de Riemann.

Antes de continuar, conviene hacer el siguiente convenio. Sea A un conjunto. Entonces, se denotará por A' al conjunto complemento de A.

Teorema 2.1.12. *(**Teorema de Portmanteau**). Las siguientes afirmaciones son equivalentes.*

(i) $\mathbf{P}_n \Rightarrow \mathbf{P}$.

(ii) $\langle \mathbf{P}_n, f \rangle \to \langle \mathbf{P}, f \rangle$ *para toda función uniformemente continua y acotada f.*

(iii) $\limsup_{n \to \infty} \mathbf{P}_n(F) \leq \mathbf{P}(F)$ *para todo cerrado F.*

(iv) $\liminf_{n \to \infty} \mathbf{P}_n(G) \geq \mathbf{P}(G)$ *para todo abierto G.*

(v) $\mathbf{P}_n(A) \to \mathbf{P}(A)$ *para todo conjunto de $\mathbf{P}-$continuidad A.*

Demostración. (ii)⇒(iii). Sea F un conjunto cerrado y $\epsilon > 0$. Se define $f : E \to [0, 1]$ por

$$f(x) = (1 - r(x, F)/\epsilon)^{+}.$$

Así como en (2.1.3), f es uniformemente continua y satisface

$$\mathbf{1}_F(x) \leq f(x) \leq \mathbf{1}_{F^\epsilon}(x).$$

Esto implica

$$\limsup_{n\to\infty} \mathbf{P}_n(F) \leq \limsup_{n\to\infty} \langle \mathbf{P}_n, f\rangle = \langle \mathbf{P}, f\rangle \leq \lim_{\epsilon\to 0} \mathbf{P}(F^\epsilon) = \mathbf{P}(F),$$

es decir,

$$\limsup_{n\to\infty} \mathbf{P}_n(F) \leq \mathbf{P}(F).$$

(iii)⇒(iv). Sea G un conjunto abierto. Entonces, por (iii), G' es abierto y satisface

$$\liminf_{n\to\infty} \mathbf{P}_n(G') \leq \limsup_{n\to\infty} \mathbf{P}_n(G') \leq \mathbf{P}(G').$$

Luego,

$$\liminf_{n\to\infty} \mathbf{P}_n(G) \geq \mathbf{P}(G).$$

(iv)⇒(v). Adicionalmente, se asume válido (iii). Sea $A \subset E$. Si $\mathbf{P}(\partial A) = 0$, entonces $\partial A = \overline{A} - A^\circ$ implica $\mathbf{P}(\overline{A}) = \mathbf{P}(A^\circ) = \mathbf{P}(A)$(pues $A^\circ \subset A \subset \overline{A}$). Ahora, como $\overline{A}$ es cerrado y A° es abierto, entonces, por (iii) y (iv),

$$\mathbf{P}(\overline{A}) \geq \limsup_{n\to\infty} \mathbf{P}_n(\overline{A}) \geq \limsup_{n\to\infty} \mathbf{P}_n(A)$$

y

$$\liminf_{n\to\infty} \mathbf{P}_n(A) \geq \liminf_{n\to\infty} \mathbf{P}_n(A^\circ) \geq \mathbf{P}(A^\circ).$$

Y ya que $\mathbf{P}(\overline{A}) = \mathbf{P}(A^\circ) = \mathbf{P}(A)$,

$$\limsup_{n\to\infty} \mathbf{P}_n(A) = \liminf_{n\to\infty} \mathbf{P}_n(A) = \mathbf{P}(A).$$

(v)⇒(i). Sea $f : E \to \mathbb{R}$ una función continua y acotada. Por linealidad se puede suponer que $0 \leq f \leq 1$. Sea $A_t = \{x : f(x) > t\}$. Por continuidad se observa que, para cada $t \in \mathbb{R}$, $\partial A_t \subset \{x : f(x) = t\}$. Para verlo, se procederá por contradicción. Se asume que $x \in \partial A_t = \overline{A}_t - A_t^\circ$ y $x \in \{x : f(x) < t\} \cup \{x : f(x) > t\}$. El primer caso implica que toda vecindad de x, V_x, satisface $V_x \cap A_t \neq \emptyset$ y $V_x \cap A_t' \neq \emptyset$. Por otro lado, por la continuidad de f, $\{x : f(x) < t\}$ y $\{x : f(x) > t\}$ son conjuntos abiertos, por lo que existe una vecindad de x completamente contenido en alguno de los dos, y claramente, esta vecindad no puede estar contenida en ambas. Pero esto contradice que $x \in \partial A$. Por lo tanto,

$$\partial A_t \subset \{x : f(x) = t\}.$$

Así, si $\mathbf{P}(\{x : f(x) = t\}) = 0$, $\mathbf{P}(A_t) = 0$. Luego por el teorema de convergencia acotada se tiene

$$\langle \mathbf{P}_n, f\rangle = \int_0^1 \mathbf{P}_n(A_t)\, dt \to \int_0^1 \mathbf{P}(A_t)\, dt = \langle \mathbf{P}, f\rangle.$$

■

Teorema 2.1.13. *Una condición necesaria y suficiente para que* $\mathbf{P}_n \Rightarrow \mathbf{P}$ *es que cada subsucesión de* $\{\mathbf{P}_n\}_{n\geq 1}$ *tenga una subsucesión que converge débilmente a* $\mathbf{P}$.

Antes de continuar, es necesario hacer el siguiente recordatorio. Sean $(\Omega, \mathcal{F}, \mathbf{P})$, E un espacio métrico y X una aplicación de Ω a E. Se dice que X es un **elemento aleatorio** de E si X es $\mathcal{F}|\mathcal{E}-$medible. Si $E = \mathbb{R}$, entonces X es una **variable aleatoria**; si $E = \mathbb{R}^k$, X es un **vector aleatorio**; y si $E = \mathbb{R}^\infty$, X es una **sucesión aleatoria**. La **distribución del elemento aleatorio** de X es la medida de probabilidad $\mathbf{P}X^{-1}$ en $(E, \mathcal{E})$ dada por[3]

$$\mathbf{P}X^{-1}(A) = \mathbf{P}\left(X^{-1}(A)\right) = \mathbf{P}\left(\omega \,\middle|\, X(\omega) \in A\right) = \mathbf{P}\left(X \in A\right).$$

Esto también es llamado la **ley** de X y se denota por $\mathcal{L}(X)$. Sean X_n y X son elementos aleatorios de E. Se dice que X_n **converge en distribución** a X, denotado por $X_n \Rightarrow X$, si $\mathbf{P}X_n^{-1} \Rightarrow \mathbf{P}X^{-1}$; se dice que $A \in \mathcal{E}$ es un $X-$**continuo** si A es de continuidad $\mathbf{P}X^{-1}$. Finalmente, si $\mathbf{P}$ es una medida de probabilidad en $(\mathbb{R}^\infty, \mathcal{R}^\infty)$, se define su **distribución finito-dimensional** por la medida $\mathbf{P}\pi_k^{-1}$. Con esto en mente, se tienen el Teorema 2.1.15 en términos de elementos aleatoria y el Teorema del mapeo continuo, el cual exhibe una relación entre dos espacio de medida bajo una función medible.

Teorema 2.1.14. ***(Caracterización alterna de convergencia en distribución.)*** *Los siguientes enunciados son equivalentes.*

(i) $X_n \Rightarrow X$.

(ii) $\mathbf{E}\left[f(X_n)\right] \to \mathbf{E}\left[f(X)\right]$ *para toda función acotada y uniformemente continua* f.

(iii) $\limsup_{n\to\infty} \mathbf{P}\left(X_n \in F\right) \leq \mathbf{P}\left(X \in F\right)$ *para todo cerrado* F.

(iv) $\liminf_{n\to\infty} \mathbf{P}\left(X_n \in G\right) \geq \mathbf{P}\left(X \in G\right)$ *para todo abierto* G.

(v) $\mathbf{P}\left(X_n \in A\right) \to \mathbf{P}\left(X \in A\right)$ *para todo conjunto* A *de* $X-$*continuido.*

Teorema 2.1.15. ***(Teorema del mapeo)***. *Sean* $(E, \mathcal{E}, d)$ *y* $(E', \mathcal{E}', d')$ *espacios métricos. Sean, además,* X_n *y* X *elementos aleatorios de* E, $h : E \to E'$ *una función medible y* D_h *el conjunto de puntos de discontinuidad de* h. *Si* $X_n \Rightarrow X$ *y* $\mathbf{P}\left(X \in D_h\right) = 0$, *entonces* $h(X_n) \Rightarrow h(X)$. *En otras palabras, dadas* $\mathbf{P}_n$ *y* $\mathbf{P}$ *medidas de probablidad en* E *tales que* $\mathbf{P}_n \Rightarrow \mathbf{P}$ *y* $\mathbf{P}\left(D_h\right) = 0$, *entonces* $\mathbf{P}_n h^{-1} \Rightarrow \mathbf{P}h^{-1}$.

Demostración. Si $x \in \overline{h^{-1}(F)}$, entonces $x_n \to x$ para alguna sucesión $\{x_n\}_{n\geq 1}$ tal que $h(x_n) \in F$. Si además $x \in D_h'$, $h(x) \in \overline{F}$. Por lo tanto,

$$D_h' \cap \overline{h^{-1}(F)} \subset h^{-1}(\overline{F}).$$

[3]Esta definición no debería causar confusión con la notación usada para definir convergencia débil ya que e

Ahora, se asume que F es cerrado. El teorema 2.1.12, junto con $\mathbf{P}\left(D_h\right)=0$, implica

$$\begin{aligned}
\limsup_{n\to\infty} \mathbf{P}_n\left(h^{-1}(F)\right) &\leq \limsup_{n\to\infty} \mathbf{P}_n\left(\overline{h^{-1}(F)}\right)\\
&\leq \mathbf{P}\left(\overline{h^{-1}(F)}\right)\\
&= \mathbf{P}\left(D_h' \cap \overline{h^{-1}(F)}\right)\\
&\leq \mathbf{P}\left(h^{-1}(\overline{F})\right)\\
&= \mathbf{P}\left(h^{-1}(F)\right),
\end{aligned}$$

es decir,

$$\limsup_{n\to\infty} \mathbf{P}_n\left(h^{-1}(F)\right) \leq \mathbf{P}\left(h^{-1}(F)\right).$$

O equivalentemente,

$$\limsup_{n\to\infty} \mathbf{P}_n h^{-1}(F) \leq \mathbf{P}h^{-1}(F).$$

Luego, por el teorema 2.1.12 inciso (iii) y la arbitrariedad del conjunto cerrado F, el resultado se sigue. ■

Se finaliza esta sección con un ejemplo que ilustra que convergencia finito dimensionales de medidas no es suficiente para tener convergencia débil de medidas. Para ello, se recuerda que (C,ρ) es el espacio métrico de funciones continuas en [0,1].

Ejemplo 2.1.16. *Para cada $n\in\mathbb{N}$, sea*

$$z_n(t) = nt\mathbf{1}_{[0,n^{-1}]}(t) + (2-nt)\mathbf{1}_{(n^{-1},2n^{-1}]}(t),\ \ t\in[0,1]. \tag{2.1.4}$$

Se observa que z_n converge uniformemente a $\mathbf{1}_{\{0\}}$, cuando $n\to\infty$. Sean $\mathbf{P}_n=\delta_{z_n}$ y $\mathbf{P}=\delta_0$, donde 0 representa la función 0. Entonces, para toda función real valuada continua y acotada en C, f, se tiene

$$\lim_{n\to\infty}\int_C f(x)\mathbf{P}_n(dx) = \lim_{n\to\infty} f(z_n) = f(\mathbf{1}_{\{0\}}) \neq f(0) = \int_C f(x)\mathbf{P}(dx),$$

es decir, $\mathbf{P}_n \not\Rightarrow \mathbf{P}$. Por otro lado, dado $t_1,\ldots,t_k$, para n suficientemente grande tal que $2n^{-1}$ sea menor que el mínimo de los t_i diferente de cero, entonces

$$\pi_{t_1,\ldots,t_k}(z_n) = (0,\ldots,0) = \pi_{t_1,\ldots,t_k}(0),$$

donde $(0,\ldots,0)$ es el vector cero en $\mathbb{R}^k$. Esto implica que, para tales n que cumplan la condición anterior,

$$\mathbf{P}_n\pi^{-1}_{t_1,\ldots,t_k}(B) = \mathbf{P}_n\left(\pi^{-1}_{t_1,\ldots,t_k}(B)\right) = \mathbf{P}\left(\pi^{-1}_{t_1,\ldots,t_k}(B)\right) = \mathbf{P}\pi^{-1}_{t_1,\ldots,t_k}(B)$$

para toda $B\subset\mathbb{R}^k$. En particular, la identidad anterior sigue siendo válido para conjuntos $\mathbf{P}\pi^{-1}_{t_1,\ldots,t_k}-$continuo. Así,

$$\lim_{n\to\infty}\mathbf{P}_n\pi^{-1}_{t_1,\ldots,t_k}(A) = \mathbf{P}\pi^{-1}_{t_1,\ldots,t_k}(A)$$

para toda conjunto A $\mathbf{P}\pi_{t_1,\ldots,t_k}^{-1}$*−continuo. Luego, por el Teorema de Portmanteau,*

$$\mathbf{P}_n\pi_{t_1,\ldots,t_k}^{-1} \Rightarrow \mathbf{P}\pi_{t_1,\ldots,t_k}^{-1},$$

aunque $\mathbf{P}_n \not\Rightarrow \mathbf{P}$. *En conclusión, convergencia débil de distribuciones finito-dimensionales no asegura convergencia débil; sin embargo, si adicionalmente las sucesiones de medidas de probabilidad son relativamente compactas, entonces se cumple la convergencia débil. Tal concepto se introduce en la siguiente sección.*

2.1.2. Tensión y teorema de Prokhorov

La propiedad de tensión de una medida (o familia de medidas) es de gran importancia. Un ejemplo de la importancia de este concepto radica en brindar un criterio para probar convergencia débil de procesos estocásticos. Intuitivamente, una medida es tensa si la medida de un conjunto compacto no se escapa en el sentido de que la medida del complemento de tal compacto es muy pequeño.

Definición 2.1.17. (i) *Se dice que una medida* $\mathbf{P}$ *en* E *es* ***tensa*** *si para todo* $\epsilon > 0$, *existe un conjunto compacto tal que* $\mathbf{P}(K) > 1 - \epsilon$.
(ii) ***Una familia de medidas de probabilidad*** Π *en* $(E, \mathcal{E})$ *se dice* ***tensa*** *si para todo* $\epsilon > 0$ *existe un compacto* $K_\epsilon \subset E$ *tal que*

$$\mathbf{P}(K_\epsilon) \geq 1 - \epsilon, \ \forall \ \mathbf{P} \in \Pi.$$

(iii) *Una familia de medidas de probabilidad* Π *en* $(E, \mathcal{E})$ *es llamada* ***relativamente compacta*** *si cualquier sucesión de* Π *contiene una subsucesión que converge débilmente.*

Ejemplo 2.1.18. *Sean* $\mathbf{P}_n$, $\mathbf{P}$ *medidas de probabilidad en* $(C, \mathcal{C})$. *Por el ejemplo 2.1.16, puede ocurrir que* $\mathbf{P}_n \not\Rightarrow \mathbf{P}$ *y*

$$\mathbf{P}_n\pi_{t_1,\ldots,t_k}^{-1} \Rightarrow \mathbf{P}\pi_{t_1,\ldots,t_k}^{-1}, \tag{2.1.5}$$

para toda k *y todo* $t_1, \ldots, t_k$. *Sin embargo, si adicionalmente se supone que la familia de medidas* $\{\mathbf{P}_n\}_{n\geq 1}$ *es relativamente compacta, entonces* $\mathbf{P}_n \Rightarrow \mathbf{P}$. *Por la compacidad relativa, cada sucesión* $\{\mathbf{P}_{n_i}\}_{i\geq 1}$ *contiene una subsucesión* $\{\mathbf{P}_{n_{i_m}}\}_{m\geq 1}$ *que converge débilmente a alguna medida de probabilidad* $\mathbf{Q}$. *Por lo que el Teorema 2.1.15 justifica*

$$\mathbf{P}_{n_{i_m}}\pi_{t_1,\ldots,t_k}^{-1} \Rightarrow \mathbf{Q}\pi_{t_1,\ldots,t_k}^{-1},$$

cuando $m \to \infty$, *y en consecuencia, por (2.1.5),*

$$\mathbf{P}\pi_{t_1,\ldots,t_k}^{-1} = \mathbf{Q}\pi_{t_1,\ldots,t_k}^{-1},$$

para toda k *y todo* $t_1, \ldots, t_k$. *Luego, recordando el ejemplo 2.1.7, el hecho de que* $\mathcal{C}_f$ *es clase separante permite deducir que* $\mathbf{P} = \mathbf{Q}$. *Se sigue del Teorema 2.1.13 que* $\mathbf{P}_n \Rightarrow \mathbf{P}$.

Otra consecuencia que se tiene de la definición de tensión son los siguientes dos resultados. La primera afirma que toda medida en un espacio Polaco (métrico, separable y compacto) satisface la propiedad de tensión. La segunda establece una manera de visualizar la medida de un conjunto a través de medidas de conjuntos compactos.

Proposición 2.1.19. *Si E es un espacio Polaco, entonces cada medida de probabilidad $\mathbf{P}$ en $(E, \mathcal{E})$ es tensa.*

Demostración. Sea $\epsilon > 0$. Como E es un epacio separable, entonces E contiene un subconjunto denso, denotado por A. Así, para cada $k \in \mathbb{N}$, se pueden tomar bolas abiertas de radio $1/k$ y centro en A y conseguir una sucesión de abiertos $A_{k,1}, A_{k,2}, \ldots$ que cubran a E[4]. Ahora, ya que $\bigcup_i A_{k,i} \uparrow E$, se puede encontrar, para cada $k \in \mathbb{N}$, n_k tal que

$$\mathbf{P}\left(\bigcup_{i \leq n_k} A_{k,i}\right) > 1 - \epsilon/2^{k+1},$$

esto es,

$$\mathbf{P}\left(\left(\bigcup_{i \leq n_k} A_{k,i}\right)'\right) < \epsilon/2^{k+1}.$$

Lo anterior permite deducir

$$\mathbf{P}\left(\bigcup_{k \in \mathbb{N}}\left(\bigcup_{i \leq n_k} A_{k,i}\right)'\right) < \sum_{k \geq 1} \epsilon/2^{k+1} = \epsilon/2 < \epsilon.$$

En otras palabras, se cumple que

$$\mathbf{P}\left(\bigcap_{k \in \mathbb{N}}\left(\bigcup_{i \leq n_k} A_{k,i}\right)\right) > 1 - \epsilon.$$

Más aún,

$$\mathbf{P}\left(\overline{H}\right) \geq \mathbf{P}\left(\bigcap_{k \in \mathbb{N}}\left(\bigcup_{i \leq n_k} A_{k,i}\right)\right) > 1 - \epsilon.$$

Para terminar la prueba, sólo falta verificar que la cerradura del conjunto

$$H := \bigcap_{k \in \mathbb{N}}\left(\bigcup_{i \leq n_k} A_{k,i}\right)$$

es compacto. Para ello, es suficiente demostrar que $\overline{H}$ es completo y totalmente acotado[5]. Como $\overline{H}$ es cerrado en un espacio métrico completo, $\overline{H}$ es completo. Por otro lado, ya que

$$H \subset \bigcup_{i \leq n_{k_0}} A_{k_0,i},$$

es decir, H puede ser cubierto por un número finito(n_{k_0} de ellos) de bolas de radio $1/k$. Como consecuencia, se obtiene que $\overline{H}$ es totalmente acotada. El resultado se sigue. ■

[4]Si $x \in E$, entonces existe una sucesión $\{a_n\} \subset A$ tal que $a_n \to x$, cuando $n \to \infty$. Esto implica, para una $k \in \mathbb{N}$ fija y algún $n_0 \in \mathbb{N}$, la existencia de un bola de radio $1/k$ y centro a_{n_0} el cual contiene a x.

[5]Sea X un espacio métrico. Entonces las siguientes afirmaciones son equivalentes: (i) X es compacto; (ii) X es completo y totalmente acotado.

Teorema 2.1.20. *Una medida de probabilidad* $\mathbf{P}$ *es tensa en* E *si, y sólo si,*

$$\mathbf{P}(A) = \sup\{\mathbf{P}(K) : K \subset A, K \text{ compacto}\}$$

para toda $A \in \mathcal{E}$.

Demostración. Sea $\mathbf{P}$ una medida tensa en E. Entonces, dado $\epsilon > 0$, existe un compacto K, un cerrado F y un abierto G tales que $\mathbf{P}(K) > 1-\epsilon/2$, $F \subset A \subset G$ y $\mathbf{P}(F - G) < \epsilon/2$ para toda $A \in \mathcal{E}$. Se nota que

$$\mathbf{P}(A - F) = \mathbf{P}(A \cap F') \leq \mathbf{P}(G \cap F') < \epsilon/2.$$

También se nota que

$$A - (F \cap K) = (A \cap F') \cup (A \cap K') \subseteq (A - F) \cup K'.$$

De lo anterior se obtiene

$$\mathbf{P}(A - (F \cap K)) \leq \mathbf{P}(A - F) + \mathbf{P}(K') < \epsilon/2 + \epsilon/2 = \epsilon.$$

Esto implica

$$\mathbf{P}(A) \leq \mathbf{P}(K \cap F) + \epsilon.$$

Ahora, debido al hecho de que $K \cap F$ es compacto (pues K es compacto y F es cerrado) y $K \cap F \subset A$,

$$\mathbf{P}(A) \leq \mathbf{P}(K \cap F) + \epsilon \leq \sup_{K \subset A} \mathbf{P}(K) + \epsilon.$$

Por lo que

$$\mathbf{P}(A) \leq \sup_{K \subset A} \mathbf{P}(K).$$

Por otro lado, para cualquier compacto K que satisfaga $K \subset A$, se tiene, por la monotonía de la medida,

$$\mathbf{P}(A) \geq \sup_{K \subset A} \mathbf{P}(K).$$

Por lo tanto

$$\mathbf{P}(A) = \sup_{K \subset A} \mathbf{P}(K).$$

Recíprocamente, si se cumple la última igualdad, entonces, tomando $A = E$, se cumple la propiedad de tensión[6]. ∎

Ejemplo 2.1.21. *Retomando la notación dada en los ejemplos 2.1.5-2.1.7, el Teorema 2.1.19 justifica que toda medida definida en los siguientes espacios son tensas:* $(\mathbb{Q}_p, |\cdot|_p)$, $\left(\mathbb{R}^k, |\cdot|\right)$, $\left(\mathbb{R}^\infty, \overline{d}\right)$ *y* (C, ρ).

[6]Basta con aplicar la definición del supremo de la siguiente manera: dado $\epsilon > 0$, existe un compacto K tal que $\mathbf{P}(K) > 1 - \epsilon$.

En el ejemplo 2.1.18 se resalta la importancia que desempeña la propiedad de compacidad relativa para probar convergencia débil. Existe otra característica que también puede conducir a convergencia débil, a saber, la propiedad de tensión. Resulta que se puede conocer cuándo una familia de medidas de probabilidad es relativamente compacta a partir de saber si la familia es tensa. Recíprocamente, agregando un par de hipótesis, se puede conocer cuándo una familia de medidas es tensa a partir de la compacidad relativa. El siguiente resultado se debe a Prohorov y proporciona un criterio para determinar si una familia de medidas de probabilidad es tensa o relativamente compacta. La prueba puede ser consultada en [4].

Teorema 2.1.22. *(**De Prohorov**) Sea (E,d) un espacio métrico y Π una familia de medidas en E.*

(i) *Si Π es tensa, entonces Π es relativamente compacta.*

(ii) *Si E es separable y completo, y Π es relativamente compacta, entonces es tensa.*

Una consecuencia inmediata del teorema anterior, es el primer criterio para probar convergencia débil de medidas.

Corolario 2.1.23. *Si $\{\mathbf{P}_n\}_{n\geq 1}$ es tensa, y si cada subsucesión que converge debilmente a $\mathbf{P}$, entonces la suceseión $\{\mathbf{P}_n\}_{n\geq 1}$ converge débilmente a $\mathbf{P}$.*

Ahora, se llevará a cabo un breve estudio del espacio de medidas de probabilidad definidas en E. Sea $\mathcal{P}(E)$ el espacio de medidas de probabilidad definidas en E. Se definirá una métrica π en $\mathcal{P}(E)$ con la propiedad de que una sucesión de medidas de probabilidad converge con respecto a π si, y sólo si, ésta converge débilmente.

Definición 2.1.24. *Sea $\mathcal{P}(E)$ el espacio de medidas de probabilidad en $(E,\mathcal{E})$. La **métrica de Prohorov** $\pi(\mathbf{P},\mathbf{Q})$ entre $\mathbf{P}$ y $\mathbf{Q}$ se define como el ínfimo sobre ϵ positivos para los cuales*

$$\mathbf{P}(A) \leq \mathbf{Q}(A^{\epsilon}) + \epsilon, \quad \mathbf{Q}(A) \leq \mathbf{P}(A^{\epsilon}) + \epsilon,$$

para toda $A \in \mathcal{E}$.

Proposición 2.1.25. *La distancia de Prohorov es una métrica en $\mathcal{P}(E)$.*

Demostración. Sean $\mathbf{P},\mathbf{Q},\mathbf{R} \in \mathcal{P}(E)$.

- Es claro que $\pi(\mathbf{P},\mathbf{Q}) = \pi(\mathbf{Q},\mathbf{P})$ y $\pi(\mathbf{P},\mathbf{P}) = 0$.
- Sea $F \subset E$ cerrado. Si $\pi(\mathbf{P},\mathbf{Q}) = 0$, entonces, para todo $\epsilon > 0$ (pues el ínfimo que satisface las condiciones de la definición de la métrica de Prohorov es 0), $\mathbf{P}(F) \leq \mathbf{Q}(F^{\epsilon}) + \epsilon$, y ya que F es cerrado, haciendo $\epsilon \to 0$ se consigue $\mathbf{P}(F) \leq \mathbf{Q}(F)$; de manera simétrica se obtiene $\mathbf{Q}(F) \leq \mathbf{P}(F)$. Luego, del Teorema 2.1.1 se tiene $\mathbf{P} = \mathbf{Q}$.
- Sean $A \in \mathcal{E}$. Si $\pi(\mathbf{P},\mathbf{Q}) < \epsilon_1$ y $\pi(\mathbf{Q},\mathbf{R}) < \epsilon_2$, entonces,

$$\mathbf{P}(A) \leq \mathbf{Q}(A^{\epsilon_1}) + \epsilon_1 \leq \mathbf{R}((\mathbf{A}^{\epsilon_1})^{\epsilon_2}) + \epsilon_1 + \epsilon_2 \leq \mathbf{R}\left(\mathbf{A}^{\epsilon_1+\epsilon_2}\right) + \epsilon_1 + \epsilon_2.$$

  Así, usando la relación simétrica, se tiene que $\pi(\mathbf{P},\mathbf{R}) < \epsilon_1 + \epsilon_2$. Por lo tanto

$$\pi(\mathbf{P},\mathbf{R}) \leq \pi(\mathbf{P},\mathbf{Q}) + \pi(\mathbf{Q},\mathbf{R}).$$

■

Teorema 2.1.26. *Si (E, d) es un espacio Polaco, entonces convergencia débil es equivalente a $\pi-$convergencia, $\mathcal{P}(E)$ es separable y completo, y una colección $\Pi \subset \mathcal{P}(E)$ es relativamente compacto si, y sólo si, su $\pi-$cerradura es $\pi-$compacto.*

Demostración. Para fines de este trabajo, interesa revelar el conjunto denso en $\mathcal{P}(E)$, por lo que sólo se probará la afirmación: si E es separable, entonces, $\mathcal{P}(E)$ es separable. La demostración completa puede verse en [4, 8]. Como E es separable, entonces E tiene una base numerable. Sea $\epsilon > 0$ y $\{A_i\}_{i\geq 1}$ una sucesión de conjuntos en E tales que A_i converge de manera creciente[7] a E y cada A_i tiene diámetro menor que ϵ, es decir,

$$\operatorname{diam}(A_i) := \sup_{x,y\in A} d(x,y) < \epsilon$$

para todo i. Para cada A_i distinto del vacío, se toma $x_i \in A_i$. Se denota por Π_ϵ el conjunto numerable conformado por medidas de probabilidad de la forma $\sum_{i\leq k} r_i\delta_{x_i}$ para algún k y números racionales $r_1, \ldots, r_k$ Ahora, dado $\mathbf{P} \in \mathcal{P}(E)$, la condición de que la sucesión conformado por los A_i converge de manera creciente a E, permite elegir $k \in \mathbb{N}$ tal que

$$\mathbf{P}\left(\bigcup_{i>k} A_i\right) < \epsilon, \tag{2.1.6}$$

y $r_1, \ldots, r_k$ racionales tales que $\sum_{i\leq k} r_i = 1$ y

$$\sum_{i\leq k} |r_i - \mathbf{P}(A_i)| < \epsilon.$$

Sea $Q = \sum_{i\leq k} r_i\delta_{x_i}$. Dado $A \subset E$, sea

$$I := \{i \leq k \,|\, A_i \subset A\}.$$

Si $A_0 = \bigcup_{i\in I} A_i$, entonces, por (2.1.6),

$$\mathbf{P}(A) \leq \mathbf{P}(A_0) + \mathbf{P}\left(\bigcup_{i>k} A_i\right) < \mathbf{P}(A_0) + \epsilon \leq \sum_{i\in I} \mathbf{P}(A_i) + \epsilon.$$

Más aún, por la elección de los racionales,

$$\mathbf{P}(A) \leq \sum_{i\in I} \mathbf{P}(A_i) + \epsilon \leq \sum_{i\in I} r_i + 2\epsilon \leq \mathbf{Q}(A_0) + 2\epsilon \leq \mathbf{Q}(A^\epsilon) + 2\epsilon,$$

lo cual implica que $\pi(\mathbf{P}, \mathbf{Q}) < 2\epsilon$.

■

[7] Para ver de manera más precisa la forma de los conjuntos que conforman la sucesión creciente, ver prueba de la Proposición 2.1.19.

Definición 2.1.27. *Sean* $\mathbf{P}, \mathbf{Q} \in \mathcal{P}(E)$. *Se define* $\mathcal{M}(\mathbf{P}, \mathbf{Q})$ *el conjunto de todas la medidas* μ *sobre* $\mathcal{E} \times \mathcal{E}$ *con* ***marginales*** $\mathbf{P}$ *y* $\mathbf{Q}$, *es decir,* $\mu(A \times E) = \mathbf{P}(A)$ *y* $\mu(E \times A) = \mathbf{Q}(A)$ *para toda* $A \in \mathcal{B}(E)$.

Lema 2.1.28. *Se cumple la siguiente relación*

$$\pi(\mathbf{P}, \mathbf{Q}) \leq \inf_{\mu \in \mathcal{M}(\mathbf{P}, \mathbf{Q})} \inf \left\{ \epsilon > 0 \,\middle|\, \mu \left\{ (x, y) \,\middle|\, d(x, y) \geq \epsilon \right\} \leq \epsilon \right\}.$$

Demostración. Si para algún $\epsilon > 0$ y $\mu \in \mathcal{M}(\mathbf{P}, \mathbf{Q})$ se tiene

$$\mu \left\{ (x, y) \,\middle|\, d(x, y) \geq \epsilon \right\} \leq \epsilon,$$

entonces, para todo conjunto cerrado F de E,

$$\begin{aligned}
\mathbf{P}(F) &= \mu(F \times E) \\
&\leq \mu\big((F \times E) \cap \left\{(x, y) \,\middle|\, d(x, y) < \epsilon\right\}\big) + \mu\big((x, y) \,\big|\, d(x, y) \geq \epsilon\big) \\
&\leq \mu(E \times F^{\epsilon}) + \epsilon \\
&= \mathbf{Q}(F^{\epsilon}) + \epsilon.
\end{aligned}$$

El resultado se sigue. ■

Proposición 2.1.29. *Sea* (E, d) *espacio separable. Sean* X_n, $n = 1, 2, \ldots$, *y* X *elementos aleatorios de* E *definidos en el mismo espacio de probabilidad con distribuciones* $\mathbf{P}_n$, $n = 1, 2, \ldots$, *y* $\mathbf{P}$, *respectivamente. Si* $\lim_{n\to\infty} d(X_n, X) = 0$ *en probabilidad, entonces* $\mathbf{P}_n \Rightarrow \mathbf{P}$.

Demostración. Para $n = 1, 2, \ldots$, sea μ_n la distribución de X_n y μ la distribución de X. Entonces

$$\lim_{n\to\infty} \mu_n \big((x, y) \,\big|\, d(x, y) \geq \epsilon\big) = \lim_{n\to\infty} \widehat{\mathbf{P}}(d(X_n, X) \geq \epsilon) = 0,$$

donde $\widehat{\mathbf{P}}$ es la medida de probabilidad en el espacio de medida donde esta definido X_n y X. Luego, por el Lema 2.1.28,

$$\lim_{n\to\infty} d(X_n, X) = 0.$$

Y ya que E es separable, el resultado se sigue del Teorema 2.1.26. ■

Antes de finalizar la sección, se enuncia el Teorema de representación de Skorohod sin prueba. Para una demostración, el lector puede consultar [4].

Teorema 2.1.30. ***(De representación de Skorokhod.)*** *Sea* (E, d) *un espacio métrico separable y sean* $\mathbf{P}$ *y* $\mathbf{P}_n$, $n = 1, 2, \ldots$, *medidas de probabilidad en* E *tales que* $\mathbf{P}_n \Rightarrow \mathbf{P}$. *Entonces, existe un espacio de probabilidad* $(\Omega, \mathcal{F}, \widehat{\mathbf{P}})$ *en el cual están definidos elementos aleatorios de* E, X *y* X_n, $n = 1, 2, \ldots$, *con distribuciones* $\mathbf{P}$ *y* $\mathbf{P}_n$, $n = 1, 2, \ldots$, *respectivamente, y tales que* X_n *converge, casi seguramente, a* X *cuando* $n \to \infty$.

2.2. Espacio $\mathcal{D}([0,\infty),E)$

Estudiar convergencia débil de procesos estocásticos requiere del análisis del espacio métrico donde las medidas de probabilidad, inducidas por los procesos estocásticos, esten definidos (más adelante se recordará la definición de medida inducida por un proceso estocástico). El espacio que se considera es el espacio de funciones definidas en $[0,\infty)$ y valores en un espacio métrico, que son continuas por la derecha y tienen límite por la izquierda. En adelante, (E,r) denota un espacio métrico y q denota la métrica $r\wedge 1$. Sea

$$\mathcal{D}([0,\infty),E)=\left\{x:[0,\infty)\to E \,\middle|\, \lim_{s\downarrow t}x(s)=x(t), \lim_{s\uparrow t}x(s)\equiv x(t-) \text{ existen}\right\},$$

el espacio de funciones *càdlàg* (del francés *continue à droite, limite à gauche*) con dominio $[0,\infty)$ y valores en E. Una propiedad interesante que cumple estas funciones es la siguiente.

Lema 2.2.1. *Si $x\in\mathcal{D}([0,\infty),E)$, entonces x tiene a lo más una cantidad numerable de puntos de discontinuidad.*

Demostración. Para $n=1,2,\ldots$, sea $A_n=\big\{t>0 \mid r\big(x(t),x(t-)\big)>1/n\big\}$. Se afirma que los conjuntos A_n son a lo más numerables. Para verlo, primero se probará que los conjuntos A_n no tienen puntos de acumulación en $[0,\infty)$. Sea n fijo y $\{t_m\}_{m\geq 1}$ una sucesión en A_n tal que

$$t_m\uparrow t \quad \text{o} \quad t_m\downarrow t,$$

cuando $m\to\infty$, para alguna $t\in[0,\infty)$. Entonces,

$$\lim_{m\to\infty}x(t_m)=x(t-)=\lim_{m\to\infty}x(t_m-)$$

o

$$\lim_{m\to\infty}x(t_m)=x(t)=\lim_{m\to\infty}x(t_m-),$$

y así,

$$r\left(x(t_m),x(t_m-)\right)<\frac{1}{n}$$

para m suficientemente grande, contradiciendo el hecho de que $t_m\in A_n$ para toda $m\geq 1$. Por lo tanto, A_n no tiene puntos de acumulación en $[0,\infty)$. Ahora, ya que, para cada $T>0$, toda sucesión en $[0,T]$ tiene una subsucesión convergente, $[0,T]\cap A_n$ consta de un número finito de puntos (de lo contrario, existiría una sucesión $\{t_k\}_{k\geq 1}$ en A_n convergente), y en consecuencia, A_n es un conjunto numerable. Luego, $\cup_{n=1}^{\infty}A_n$, el conjunto de todas las discontinuidades de x, es numerable. ■

Corolario 2.2.2. *Para cada $n\in\mathbb{N}$, sea A_n como en el Lema 2.2.1. Entonces, para cada $T>0$, $[0,T]\cap A_n$ es un conjunto finito.*

Debido al problema de definir una métrica que permita comparar dos funciones *càdlàg* con puntos de discontinuidad diferentes, se consideran los siguientes conjuntos

$$\Lambda'=\left\{\lambda:[0,\infty)\to[0,\infty) \,\middle|\, \lambda \text{ es estrictamente creciente}\right\}$$

y

$$\Lambda = \left\{ \lambda \in \Lambda' \;\middle|\; \lambda \text{ es Lipschitz continua y } \gamma(\lambda) < \infty \right\},$$

donde

$$\gamma(\lambda) := \sup_{s>t\geq 0} \left| \log \frac{\lambda(s)-\lambda(t)}{s-t} \right|.$$

En particular, para cada $\lambda \in \Lambda'$, $\lambda(0) = 0$, $\lim_{t\to\infty} \lambda(t) = \infty$ y λ es continua. En la literatura, el término $\gamma(\lambda)$ se le conoce la dilatación de λ. Para $x, y \in \mathcal{D}([0,\infty), E)$, se define la siguiente aplicación

$$d(x,y) = \inf_{\lambda\in\Lambda} \left[\gamma(\lambda) \vee \int_0^\infty \exp\{-u\} d'(x,y,\lambda,u) du \right] \tag{2.2.1a}$$

con

$$d'(x,y,\lambda,u) = \sup_{t\geq 0} q\Big(x(t\wedge u), y(\lambda(t)\wedge u) \Big). \tag{2.2.1b}$$

Lema 2.2.3. *La aplicación dilatación $\gamma : \Lambda \to \mathbb{R}^+$ satisface las siguientes propiedades.*

(i) *Para toda $\lambda \in \Lambda$, $\gamma(\lambda) = \gamma\left(\lambda^{-1}\right)$.*

(ii) *Si $\lambda_1, \lambda_2 \in \Lambda$, entonces $\lambda_1 \circ \lambda_2 \in \Lambda$ y*

$$\gamma(\lambda_1 \circ \lambda_2) \leq \gamma(\lambda_1) + \gamma(\lambda_2).$$

(iii) *Si $\{\lambda_n\}_{n\geq 1} \subset \Lambda$ con $\gamma(\lambda_n) \to 0$, cuando $n \to \infty$, entonces, para toda $T \geq 0$,*

$$\lim_{n\to\infty} \sup_{t\in[0,T]} |\lambda_n(t) - t| = 0.$$

Demostración. **(i)** Para toda $\lambda \in \Lambda$ se tiene, tomando $s' = \lambda(s)$ y $t' = \lambda(t)$ para $s > t \geq 0$, el hecho de que la función inversa de λ sea estrictamente creciente implica

$$\gamma(\lambda) = \sup_{s>t\geq 0} \left| \log \frac{\lambda(s)-\lambda(t)}{s-t} \right| = \sup_{s'>t'\geq 0} \left| -\log \frac{\lambda^{-1}(s')-\lambda^{-1}(t')}{s'-t'} \right| = \gamma(\lambda^{-1}).$$

(ii) Sean $\lambda_1, \lambda_2 \in \Lambda$. Entonces, $\lambda_1 \circ \lambda_2$ es Lipschitz continua y además, usando la notación dada en el primer apartado,

$$\begin{aligned}
\gamma(\lambda_1 \circ \lambda_2) &= \sup_{s>t\geq 0} \left| \log \frac{\lambda_1 \circ \lambda_2(s) - \lambda_1 \circ \lambda_2(t)}{s-t} \right| \\
&= \sup_{s>t\geq 0} \left| \log \frac{\lambda_1\big(\lambda_2(s)\big) - \lambda_1\big(\lambda_2(t)\big)}{\lambda_2(s)-\lambda_2(t)} + \log \frac{\lambda_2(s)-\lambda_2(t)}{t-s} \right| \\
&\leq \sup_{s'>t'\geq 0} \left| \log \frac{\lambda_1\big(s'\big) - \lambda_1\big(t'\big)}{s'-t'} \right| + \sup_{s>t\geq 0} \left| \log \frac{\lambda_2(s)-\lambda_2(t)}{t-s} \right| \\
&\leq \gamma(\lambda_1) + \gamma(\lambda_2).
\end{aligned}$$

(iii) Primero, se observa que para $\lambda \in \Lambda$ se cumple

$$\begin{aligned}\gamma(\lambda) &\geq 1 - \exp\{-\gamma(\lambda)\} \\ &= \sup_{0\leq t<s}\left(1 - \exp\left\{-\left|\log\frac{\lambda(s)-\lambda(t)}{s-t}\right|\right\}\right) \\ &= \sup_{0\leq t<s}\left(1 - \frac{\lambda(s)-\lambda(t)}{s-t} \wedge \frac{s-t}{\lambda(s)-\lambda(t)}\right) \\ &= \sup_{0\leq t<s}\left(1 - \frac{\lambda(s)-\lambda(t)}{s-t}\right).\end{aligned}$$

Así, si $\{\lambda_n\}_{n\geq 1} \subset \Lambda$ es tal que $\gamma(\lambda_n) \to 0$, cuando $n \to \infty$, entonces

$$\sup_{0\leq t<s}\frac{\lambda_n(s)-\lambda_n(t)}{s-t} \to 1, \quad n \to \infty.$$

En particular, para toda $T \geq 0$,

$$\lim_{n\to\infty}\sup_{t\in[0,T]}|\lambda_n(t) - t| = 0.$$

■

Más adelante, en la Proposición 2.2.6 se probará que $(\mathcal{D}([0,\infty),E),d)$ es un espacio métrico. Para ello, primero se demostrarán algunos resultados preliminares.

Lema 2.2.4. *Sean $\{x_n\}_{n\geq 1}, \{y_n\}_{n\geq 1} \subset \mathcal{D}([0,\infty),E)$. Entonces, $\lim_{n\to\infty} d(x_n,y_n) = 0$ si, y sólo si, existe una sucesión $\{\lambda_n\}_{n\geq 1} \subset \Lambda$ tal que $\lim_{n\to\infty}\gamma(\lambda_n) = 0$ y*

$$\lim_{n\to\infty} m\left(u \in [0,T] \,\middle|\, d'(x_n,y_n,\lambda_n,u) \geq \epsilon\right) = 0 \tag{2.2.2}$$

para toda $\epsilon > 0$ y toda $T > 0$, donde m denota la medida de Lebesgue.

Demostración. La prueba hace uso del teorema de convergencia de Vitali (ver Teorema 7.13 de [3]) y del siguiente resultado. Sea f una función acotada en un intervalo acotado de la forma $[a,b]$. Entonces, f es Riemann intregrable en $[a,b]$ si, y sólo si, f es continuo m-casi en todas partes. En este caso, f es Lebesgue integrable en $[a,b]$ y la integral de Lebesgue de f coincide con la integral de Riemann de f (ver Teorema 2.4.1 de [2]).
Si $\lim_{n\to\infty} d(x_n,y_n) = 0$, por definición y propiedad de ínfimo, existe, para todo $\epsilon > 0$, una sucesión $\{\lambda_n\}_{n\geq 1}$ tal que $\gamma(\lambda_n) < \epsilon$ y

$$\int_0^\infty \exp\{-u\}d'(x_n,y_n,\lambda_n,u)du < \epsilon. \tag{2.2.3}$$

para n suficientemente grande. En consecuencia, se tiene $\lim_{n\to\infty}\gamma(\lambda_n) = 0$ y que

$$f_n(u) := \exp\{-u\}d'(x_n,y_n,\lambda_n,u)$$

es acotado y es Rieman integrable en todo intervalo acotado de la forma $[a,b]$ para n suficientemente grande. Aplicando el resultado dado al principio de la prueba, f_n es continua

m-casi en todas partes y la integral de Lebesgue de f_n coincide con la integral de Riemann de f_n para n suficientemente grande, es decir,

$$\int_{[a,b]} \exp\{-u\}d'(x_n, y_n, \lambda_n, u)dm = \int_{[a,b]} \exp\{-u\}d'(x_n, y_n, \lambda_n, u)du.$$

Así, por (2.2.3), para todo $\epsilon > 0$, $f_n < \epsilon$ $m-$casi en todas partes en todo intervalo acotado de la forma $[a, b]$ y para n suficientemente grande. Esto implica, para todo $\epsilon > 0$, $d'(x_n, y_n, \lambda_n, u) < \epsilon$ $m-$casi en todas partes en todo intervalo acotado de la forma $[a, b]$ y para n suficientemente grande, y así,

$$m\left(u \in [0, T] \,\middle|\, d'(x_n, y_n, \lambda_n, u) \geq \epsilon\right) = 0 \tag{2.2.4}$$

para toda $\epsilon > 0$, $T > 0$ y n suficientemente grande. Se sigue (2.2.2). El recíproco se obtiene aplicando el teorema de Vitali a las funciones real valuadas f_n dadas por

$$f_n(u) = \exp\{-u\}d'(x_n, y_n, \lambda_n, u), \quad n \geq 1,$$

donde $x, x_n \in \mathcal{D}([0, \infty), E)$, $n \geq 1$ y $u \in \mathbb{R}^+$. ■

Lema 2.2.5. *Sean $x, x_n \in \mathcal{D}([0, \infty), E)$ para $n \in \mathbb{N}$. Entonces, $\lim_{n\to\infty} d(x_n, x) = 0$ si, y sólo si, existe una sucesión $\{\lambda_n\}_{n\geq 1} \subset \Lambda$ tal que $\lim\limits_{n\to\infty} \gamma(\lambda_n) = 0$ y*

$$\lim_{n\to\infty} d'(x_n, x, \lambda_n, u) = 0 \tag{2.2.5}$$

para todo punto de continuidad u de x.

Demostración. Primero, se asume la existencia $\{\lambda_n\}_{n\geq 1} \subset \Lambda$ tal que $\lim\limits_{n\to\infty} \gamma(\lambda_n) = 0$ y (2.2.5). Sea T un real positivo. Se denota por $C_T(x)$ el conjunto de puntos de discontinuidad de x en el intervalo $[0, T]$. Entonces, por el Lema de Fatou y (2.2.5),

$$\begin{aligned}
\lim_{n\to\infty} m\left(u \in [0, T] \,\middle|\, d'(x_n, y_n, \lambda_n, u) \geq \epsilon\right) &\leq \limsup_{n\to\infty} m\left(u \in [0, T] \,\middle|\, d'(x_n, x, \lambda, u) \geq \epsilon\right) \\
&\leq m\left(u \in [0, T] \,\middle|\, \limsup_{n\to\infty} d'(x_n, x, \lambda, u) \geq \epsilon\right) \\
&\leq m\left(C_T(x)\right),
\end{aligned}$$

siendo este último término igual a cero ya que el conjunto de puntos de discontinidades de x es numerable(Lema 2.2.1). Luego, por el Lema 2.2.4, $\lim_{n\to\infty} d(x_n, x) = 0$.

Recíprocamente, se supone válido $\lim_{n\to\infty} d(x_n, x) = 0$ y u un real no negativo. Sea $\{\lambda_n\}_{n\geq 1} \subset \Lambda$ la sucesión que satisface el lema 2.2.4 con $y_n = x$ para toda $n \in \mathbb{N}$, esto es, $\lim\limits_{n\to\infty} \gamma(\lambda_n) = 0$ y

$$\lim_{n\to\infty} m\left(u \in [0, T] \,\middle|\, d'(x_n, x, \lambda_n, u) \geq \epsilon\right) = 0,$$

para toda $\epsilon > 0$ y $T > 0$, donde m denota la medida de Lebesgue. En particular, existe una sucesión $\{N_k\}_{k\geq 1}$ tal que para toda $n \geq N_k$

$$m\left(v \in (u, u+1] \,\middle|\, d'(x_n, x, \lambda_n, v) < \frac{1}{k}\right) > 0, \tag{2.2.6}$$

y con esto, para cada $N_k \leq n < N_{k+1}$, existe un $u_n \in (u, u+1]$ tal que

$$d'(x_n, x, \lambda_n, u_n) < \frac{1}{k}$$

Así, eligiendo valores $u_n \in (u, u+1]$ para $n \geq N_1$, se obtiene

$$\lim_{n\to\infty} \sup_{t\geq 0} q\Big(x_n(t\wedge u_n), x(\lambda_n(t)\wedge u_n)\Big) = \lim_{n\to\infty} d'\big(x_n, x, \lambda_n, u_n\big) = 0. \tag{2.2.7}$$

Ahora, para cada $n \in \mathbb{N}$,

$$\begin{aligned} d'\big(x_n, x, \lambda_n, u\big) &= \sup_{t\geq 0} q\Big(x_n(t\wedge u), x(\lambda_n(t)\wedge u)\Big) \\ &\leq \sup_{t\geq 0} q\Big(x_n(t\wedge u), x(\lambda_n(t\wedge u)\wedge u_n)\Big) \\ &\quad + \sup_{t\geq 0} q\Big(x(\lambda_n(t\wedge u)\wedge u_n), x(\lambda_n(t)\wedge u)\Big). \end{aligned}$$

Pero, para cada n,

$$\begin{aligned} \sup_{t\geq 0} q\Big(x(\lambda_n(t\wedge u)\wedge u_n), x(\lambda_n(t)\wedge u)\Big) &= \sup_{0\leq t\leq u} q\Big(x(\lambda_n(t\wedge u)\wedge u_n), x(\lambda_n(t)\wedge u)\Big) \\ &\quad \vee \sup_{t> u} q\Big(x(\lambda_n(t\wedge u)\wedge u_n), x(\lambda_n(t)\wedge u)\Big) \\ &= \sup_{0\leq t\leq u} q\Big(x(\lambda_n(t)\wedge u_n), x(\lambda_n(t)\wedge u)\Big) \\ &\quad \vee \sup_{t> u} q\Big(x(\lambda_n(u)\wedge u_n), x(\lambda_n(t)\wedge u)\Big) \\ &= \sup_{u\leq s\leq \lambda_n(u)\vee u} q\Big(x(s), x(u)\Big) \\ &\quad \vee \sup_{\lambda_n(u)\wedge u< s\leq u} q\Big(x(\lambda_n(u)\wedge u_n), x(s)\Big), \end{aligned}$$

donde la última igualdad se obtiene estableciendo $s = \lambda_n(t)\wedge u_n$ en el primer término, y $s = \lambda_n(t)\wedge u$ en el segundo término. Así, para cada $n \in \mathbb{N}$,

$$\begin{aligned} d'(x_n, x, \lambda_n, u) \leq & \sup_{0\leq t\leq u} q\Big(x_n(t\wedge u_n), x(\lambda_n(t)\wedge u_n)\Big) \\ & + \sup_{u\leq s\leq \lambda_n(u)\vee u} q\Big(x(s), x(u)\Big) \\ & \vee \sup_{\lambda_n(u)\wedge u< s\leq u} q\Big(x(\lambda_n(u)\wedge u_n), x(s)\Big) \end{aligned} \tag{2.2.8}$$

Luego, usando el Lema 2.2.3 inciso (iii), (2.2.7) y u un punto de continuidad de x, la última desigualdad implica

$$\lim_{n\to\infty} d'(x_n, x, \lambda_n, u) = 0.$$

■

Proposición 2.2.6. *El espacio* $(\mathcal{D}([0,\infty),E),d)$ *es un espacio métrico.*

Demostración.

- Sea $x,y\in\mathcal{D}([0,\infty),E)$. Se probará que $d(x,y)=0$ si, y sólo si, $x=y$. Si $x=y$, entonces, tomando λ_n la función identidad para toda $n\in\mathbb{N}$, $d(y,x)=0$.

  Ahora, se asume $d(y,x)=0$. Tomando $x_n=y$ para toda n en el Lema 2.2.5, existe una sucesión $\{\lambda_n\}_{n\geq 1}\subset\Lambda$ tal que

  $$\lim_{n\to\infty} d'(y,x,\lambda_n,u)=0$$

  para todo punto de continuidad u de x. En particular, para todo punto de continuidad u de x y $t>0$, se tiene

  $$\lim_{n\to\infty} r\big(y(t\wedge u),x(\lambda_n(t)\wedge u)\big)=0.$$

  Por lo que, eligiendo $t=u$, el Lema 2.2.3 (iii) permite deducir

  $$r\big(y(u),x(u)\big)=0.$$

  Esto es, $x(u)=y(u)$ para todos los puntos de continuidad u de x. Por lo tanto, por el Lema 2.2.1 y la continuidad por la derecha de x y y, $x=y$.

- Se observa que, para $x,y,\in\mathcal{D}([0,\infty),E)$,

  $$\sup_{t\geq 0} q\Big(x(t\wedge u),y(\lambda(t)\wedge u)\Big)=\sup_{t\geq 0} q\Big(x(\lambda(t)^{-1}\wedge u),y(t\wedge u)\Big)$$

  para toda $\lambda\in\Lambda$. Así, la reflexividad es consecuencia de esta identidad y $\gamma(\lambda)=\gamma(\lambda^{-1})$(Lema 2.2.3 (i)). Por lo tanto, $d(x,y)=d(y,x)$.

- Se verificará la desigualdad del triángulo. Sean $x,y,z\in\mathcal{D}([0,\infty),E)$. Entonces, para todo $u\geq 0$ se tiene

  $$\begin{aligned}
  \sup_{t\geq 0} q\Big(x(t\wedge u),y\big(\lambda_1\circ\lambda_2(t)\wedge u\big)\Big)&\\
  &\leq \sup_{t\geq 0} q\Big(x(t\wedge u),y\big(\lambda_2(t)\wedge u\big)\Big)\\
  &\quad+\sup_{t\geq 0} q\Big(y(\lambda_2(t)\wedge u),z\big(\lambda_1\circ\lambda_2(t)\wedge u\big)\Big)\\
  &= \sup_{t\geq 0} q\Big(x(t\wedge u),y\big(\lambda_2(t)\wedge u\big)\Big)\\
  &\quad+\sup_{t\geq 0} q\Big(y(t\wedge u),z\big(\lambda_1(t)\wedge u\big)\Big).
  \end{aligned}$$

  Por otro lado, se recuerda que el Lema 2.2.3 (ii) afirma

  $$\gamma(\lambda_1\circ\lambda_2)\leq\gamma(\lambda_1)+\gamma(\lambda_2).$$

  Combinando las desigualdades anteriores se obtiene

  $$d(x,z)\leq d(x,y)+d(y,z)$$

Por lo tanto, d es un métrica en $\mathcal{D}([0,\infty),E)$. ■

Definición 2.2.7. *La métrica d es llamada **métrica de Skorohod,** y la topólogía inducida por esta métrica es llamada **topología de Skorohod**.*

Proposición 2.2.8. *Sea $\{x_n\}_{n\geq 1} \subset \mathcal{D}([0,\infty),E)$. Entonces las siguientes afirmaciones son equivalentes.*

(i) $\lim\limits_{n\to\infty} d(x_n,x)=0.$

(ii) *Existe una sucesión $\{\lambda_n\}_{n\geq 1} \subset \Lambda$ tal que $\lim_{n\to\infty}\gamma(\lambda_n)=0$ y*

$$\lim_{n\to\infty}\sup_{0\leq t\leq T} r\big(x_n(t),x(\lambda_n(t))\big)=0 \tag{2.2.9}$$

para toda $T>0$.

(iii) *Para cada $T>0$, existe $\{\lambda_n\}_{n\geq 1} \subset \Lambda'$ (posiblemente depende de T) tal que (2.2.9) y*

$$\lim_{n\to\infty}\sup_{0\leq t\leq T}|\lambda_n(t)-t|=0.$$

se sostienen.

Demostración. (i) ⇒ (ii). Sean $\{x_n\}_{n\geq 1}, x\in\mathcal{D}([0,\infty),E)$. Por el Lema 2.2.4, existe una sucesión $\{\lambda_n\}_{n\geq 1} \subset \Lambda$ tal que $\lim\limits_{n\to\infty}\gamma(\lambda_n)=0$ y

$$\lim_{n\to\infty} m\left(u\in[0,T] \,\middle|\, d'(x_n,x,\lambda_n,u)\geq\epsilon\right)=0$$

para toda $\epsilon>0$ y $T>0$, donde m denota la medida de Lebesgue. En particular, por (2.2.6) con $u=k$, existe una sucesión $\{u_n\}_{n\geq 1}\subset(0,\infty)$ con $u_n\to\infty$, cuando $n\to\infty$, y tal que

$$\lim_{n\to\infty} d'\big(x_n,x,\lambda_n,u_n\big)=0.$$

En consecuencia,

$$\lim_{n\to\infty}\sup_{t\geq 0} r\Big(x_n(t\wedge u_n),x(\lambda_n(t)\wedge u_n)\Big)=0.$$

Ahora, dado $T\in[0,\infty)$, y tomando $u_n\geq T\vee\lambda_n(T)$ para n suficientemente grande, el límite anterior implica

$$\lim_{n\to\infty}\sup_{0\leq t\leq T} r\Big(x_n(t),x(\lambda_n(t))\Big)\leq \lim_{n\to\infty}\sup_{t\geq 0} r\Big(x_n(t\wedge u_n),x(\lambda_n(t)\wedge u_n)\Big)=0.$$

Por la arbitrariedad de T el resultado se sigue.

(ii) ⇒ (i). Sean $\{\lambda_n\}_{n\geq 1}\subset\Lambda$ una sucesión que satisface la condición (ii) y $u\in[0,\infty)$. Además, se considera una sucesión $\{u_n\}_{n\geq 1}$ tal que $u_n\to\infty$, cuando $n\to\infty$. Se recuerda, por (2.2.8),

$$\begin{aligned} d'(x_n, x, \lambda_n, u) \leq & \sup_{0 \leq t \leq u} q\Big(x_n(t \wedge u_n), x(\lambda_n(t) \wedge u_n)\Big) \\ & + \sup_{u \leq s \leq \lambda_n(u) \vee u} q\Big(x(s), x(u)\Big) \\ & \vee \sup_{\lambda_n(u) \wedge u < s \leq u} q\Big(x(\lambda_n(u) \wedge u_n), x(s)\Big) \end{aligned}$$

Se observa que, para n suficientemente grande tal que $u_n \geq \lambda_n(u) \vee u$, el primer término de la derecha de la última desigualdad satisface

$$\sup_{0 \leq t \leq u} q\Big(x_n(t \wedge u_n), x(\lambda_n(t) \wedge u_n)\Big) = \sup_{0 \leq t \leq u} q\Big(x_n(t), x(\lambda_n(t))\Big) \leq \sup_{0 \leq t \leq u} r\Big(x_n(t), x(\lambda_n(t))\Big)$$

Esto conduce, para n suficientemente grande para que $u_n \geq \lambda_n(u) \vee u$, a la siguiente desigualdad

$$\begin{aligned} d'(x_n, x, \lambda_n, u) \leq & \sup_{0 \leq t \leq u} r\Big(x_n(t), x(\lambda_n(t))\Big) \\ & + \sup_{u \leq s \leq \lambda_n(u) \vee u} q\Big(x(s), x(u)\Big) \\ & \vee \sup_{\lambda_n(u) \wedge u < s \leq u} q\Big(x(\lambda_n(u) \wedge u_n), x(s)\Big). \end{aligned}$$

Luego, para u un punto de continuidad de x, la hipótesis y la continuidad por la derecha de x implican

$$\lim_{n \to \infty} d'(x_n, x, \lambda_n, u) = 0.$$

Por lo tanto, existe una sucesión $\{\lambda_n\}_{n \geq 1} \subset \Lambda$ que satisface $\lim\limits_{n \to \infty} \gamma(\lambda_n) = 0$ y

$$\lim_{n \to \infty} d'(x_n, x, \lambda_n, u) = 0.$$

El resultado se sigue del Lema 2.2.5.

(iii) $\Rightarrow$ (ii). Sea N un entero fijo y $\{\lambda_n^N\}_n \subset \Lambda'$ que satisface las hipótesis de (iii) con $T = N$ tal que

$$\lambda_n^N(t) := \lambda_n^N(N) + t - N, \quad t \geq N.$$

Se construirá una sucesión $\{\hat{\mu}_n^N\}_n$ en Λ que cumpla con las condiciones establecidas en (ii). Se define $\tau_0^N := 0$, y para toda $k \geq 1$,

$$\tau_k^N := \inf\left\{ t > \tau_{k-1} \mid r\left(x(t), x(\tau_{k-1}^N)\right) > \frac{1}{N} \right\}$$

si $\tau_{k-1}^N < \infty$, $\tau_k^N = \infty$ si $\tau_{k-1}^N = \infty$. En otras palabras, el tiempo se particiona de tal modo que la distancia de los elementos de cada partición bajo x se encuentran a una distancia menor que $1/N$. Se observa que, debido a la continuidad por la derecha de x, la sucesión $\{\tau_n^N\}_{n \geq 1}$ es estrictamente creciente. Más aún, la existencia del límite por la izquierda de

x garantiza que la sucesión no tenga ningún punto límite (de lo contrario, la función x tendría que tender a infinito en algún tiempo finito y esto contradiría la existencia del límite por la izquierda).

Ahora, para cada $n \in \mathbb{N}$, sea

$$\mu_{k,n}^N := (\lambda_n^N)^{-1}(\tau_k^N)$$

para $k = 0, 1, \ldots$, donde $(\lambda_n^N)^{-1}(\infty) = \infty$. Se observa que, para toda $t \in [0, N]$ y toda $k \in \mathbb{N}$, $\mu_{k,n}^N \leq t < \mu_{k+1,n}^N$ implica

$$\tau_k^N \leq \lambda_n^N(t) < \tau_{k+1}^N. \tag{2.2.10}$$

También, se observa que, para toda k y n naturales, se cumple

$$0 \leq \frac{\tau_k^N - \tau_{k-1}^N}{\mu_{k,n}^N - \mu_{k-1,n}^N} \leq \frac{\tau_{k+1}^N - \tau_k^N}{\mu_{k+1,n}^N - \mu_{k,n}^N}.$$

Ahora, para cada $n \in \mathbb{N}$, se define la siguiente aplicación

$$\mu_n^N(t) := \begin{cases} \tau_k^N + \left(t - \mu_{k,n}^N\right) \dfrac{\left(\tau_{k+1}^N - \tau_k^N\right)}{\left(\mu_{k+1,n}^N - \mu_{k,n}^N\right)}, & \text{si} \quad t \in [\mu_{k,n}^N, \mu_{k+1,n}^N) \cap [0, N],\, k = 0, 1, \ldots, \\ \\ \mu_n^N(N) + t - N, & \text{si} \qquad\qquad t > N. \end{cases}$$

Sea $n \in \mathbb{N}$ dado. Primero, se observa que si

$$[\mu_{k,n}^N, \mu_{k+1,n}^N) \cap [0, N] \neq \emptyset \tag{2.2.11}$$

para alguna $k \geq 1$,

$$\tau_k^N \leq \mu_n^N(t) \leq \tau_{k+1}^N \tag{2.2.12}$$

para toda $t \in [\mu_{k,n}^N, \mu_{k+1,n}^N) \cap [0, N]$. Más aún, μ_n^N es una aplicación continua en $[\mu_{k,n}^N, \mu_{k+1,n}^N) \cap [0, N]$ por el criterio de la primera derivada. En adelante, se asumirá (2.2.11). Si $s \in [\mu_{k-1,n}^N, \mu_{k,n}^N) \cap [0, N]$ y $t \in [\mu_{k,n}^N, \mu_{k+1,n}^N) \cap [0, N]$, entonces

$$\mu_n^N(t) - \mu_n^N(s) \geq (t - s)\frac{\tau_{k+1}^N - \tau_k^N}{\mu_{k+1,n}^N - \mu_{k,n}^N} > 0.$$

Si $s \in [\mu_{k-1,n}^N, \mu_{k,n}^N) \cap [0, N]$ y $t \in [\mu_{k+1,n}^N, \mu_{k+2,n}^N) \cap [0, N]$, entonces la desigualdad anterior conduce a

$$\begin{aligned} \mu_n^N(t) - \mu_n^N(s) &= \mu_n^N(t) - \mu_n^N\left(\mu_{k,n}^N\right) + \mu_n^N\left(\mu_{k,n}^N\right) - \mu_n^N(s) \\ &\geq \left(t - \mu_k^N\right) \frac{\tau_{k+2}^N - \tau_k^N}{\mu_{k+2,n}^N - \mu_{k,n}^N} + \left(\mu_{k,n}^N - s\right) \frac{\tau_{k+1}^N - \tau_k^N}{\mu_{k+1,n}^N - \mu_{k,n}^N} \\ &\geq \left(t - \mu_{k,n}^N\right) \frac{\tau_{k+1}^N - \tau_k^N}{\mu_{k+1,n}^N - \mu_{k,n}^N} + \left(\mu_{k,n}^N - s\right) \frac{\tau_{k+1}^N - \tau_k^N}{\mu_{k+1,n}^N - \mu_{k,n}^N} \\ &\geq (t - s)\frac{\tau_{k+1}^N - \tau_k^N}{\mu_{k+1,n}^N - \mu_{k,n}^N}, > 0, \end{aligned}$$

y así,

$$\mu_n^N(t) - \mu_n^N(s) > 0.$$

El mismo resultado se obtiene si $t > N \geq s$. Este razonamiento permite concluir, para toda $n \in \mathbb{N}$, que la aplicación $\mu_n^N(t)$ es estrictamente creciente respecto a t. Además, si $t \in [\mu_{k,n}^N, \mu_{k+1,n}^N) \cap [0, N]$, entonces, tomando el convenio $\infty^{-1}\infty = 1$,

$$\mu_n^N(t)' = \frac{\tau_{k+1}^N - \tau_k^N}{\mu_{k+1,n}^N - \mu_{k,n}^N} \leq \text{máx}\left\{ \frac{\tau_{k+1}^N - \tau_k^N}{\mu_{k+1,n}^N - \mu_{k,n}^N} \,\middle|\, [\mu_{k,n}^N, \mu_{k+1,n}^N) \cap [0,N] \neq \emptyset, k = 0, 1, \ldots \right\} \vee 1.$$

para toda $n \in \mathbb{N}$. Esto implica que $\mu_n^N(t)'$ está acotada para toda $t \leq N$. Si $t > N$, entonces $\mu_n^N(t)' = 1$. Por lo tanto, $\mu_n^N(t)$ es Lipschitz continua para toda n, y en consecuencia, $\{\mu_n^K\}_{n \geq 1} \subset \Lambda$. Por otro lado, para cada $n \in \mathbb{N}$ se cumple

$$\left|\mu_n^N(t) - \mu_n^N(s)\right| = |t - s| \left|(\tau_{k+1,n}^N - \tau_{k,n}^N)(\mu_{k+1,n}^N - \mu_{k,n}^N)^{-1}\right|$$

para $s, t \in [\mu_{k,n}^N, \mu_{k+1,n}^N) \cap [0, N]$, $k = 0, 1, \ldots$, y

$$\left|\mu_n^N(t) - \mu_n^N(s)\right| = |t - s|$$

para $s, t \geq N$. Más aún,

$$\left|\log \frac{\mu_n^N(t) - \mu_n^N(s)}{t - s}\right| = \left|\log \frac{\tau_{k+1,n}^N - \tau_{k,n}^N}{\mu_{k+1,n}^N - \mu_{k,n}^N}\right| \tag{2.2.13}$$

para $s, t \in [\mu_{k,n}^N, \mu_{k+1,n}^N) \cap [0, N]$, $k = 0, 1, \ldots$, y

$$\left|\log \frac{\mu_n^N(t) - \mu_n^N(s)}{t - s}\right| = 0 \tag{2.2.14}$$

para $s, t \geq N$. Luego, usando nuevamente el convenio $\infty^{-1}\infty = 1$, (2.2.13) y (2.2.14) implican

$$\gamma\left(\mu_n^N\right) = \sup_{0 \leq s < t} \left|\log \frac{\mu_n^N(t) - \mu_n^N(s)}{t - s}\right| = \text{máx}_{\mu_{k,n}^N \leq N} \left|\log \frac{\tau_{k+1,n}^N - \tau_{k,n}^N}{\mu_{k+1,n}^N - \mu_{k,n}^N}\right|,$$

y, junto con (2.2.10), (2.2.12) y la definición τ_k,

$$\sup_{0 \leq t \leq N} r\left(x(\lambda_n^N(t))x(\mu_n^N(t))\right) \leq \frac{2}{N} \tag{2.2.15}$$

para toda $n \geq 1$. Luego, por hipótesis con $T = N$ y por (2.2.15),

$$\lim_{n \to \infty} \gamma\left(\mu_n^N\right) = \lim_{n \to \infty} \text{máx}_{\mu_{k,n}^N \leq N} \left|\log\left(\tau_{k+1}^N - \tau_{k,n}^N\right) - \log\left(\mu_{k+1,n}^N - \mu_{k,n}^N\right)\right| = 0$$

y

$$\begin{aligned}
\sup_{0 \leq t \leq N} r(x_n(t), x(\mu_n^N(t))) &\leq \sup_{0 \leq t \leq N} r(x_n(t), x(\lambda_n^N(t))) + \sup_{0 \leq t \leq N} r(x_n(\lambda_n^N(t))), x(\mu_n^N(t))) \\
&\leq \sup_{0 \leq t \leq N} r(x_n(t), x(\lambda_n^N(t))) + \frac{2}{N} \\
&\leq \frac{3}{N}.
\end{aligned}$$

para toda $n \in \mathbb{N}$. Esto asegura la elección de una subsucesión $\{n_k\}_{k\geq 1}$ tal que

$$\gamma(\mu_n^N) \leq \frac{1}{N}, \quad \text{y} \quad \sup_{0\leq t\leq N} r(x_n(t), x(\mu_n^N(t))) \leq \frac{3}{N}$$

para $n > n_N$. Entonces, para $1 \leq n < n_1$ sea $\hat{\lambda}_n$ arbitraria; mientras que para cualquier $n_N \leq n < n_{N+1}$, con $N \geq 1$, sea $\hat{\lambda}_n := \mu_n^N$. Así, la sucesión $\{\hat{\lambda}_n\}_{n\geq 1}$ satisface las propiedades enunciadas en (ii), y con esto se finaliza la prueba. ■

Ejemplo 2.2.9. *Sea $\alpha \geq 0$ y d_u la métrica uniforme dada por*

$$d_u(x,y) := \sup_{t\geq 0} |x(t) - y(t)|, \quad x, y \in \mathcal{D}([0,\infty), E).$$

Se definen $x_n := \mathbf{1}_{[0,\alpha+1/n)}$ y $x := \mathbf{1}_{[0,\alpha)}$. Entonces $x_n \not\to x$ bajo d_u (falla en $t = \alpha$), pero $x_n \to x$ bajo la métrica de Skorokhod. Para verificarlo, sólo basta aplicar la proposición anterior con la sucesión de funciones $\{\lambda_n\}_{n\geq 1}$ dada por

$$\lambda_n(t) := \begin{cases} \text{lineal} & \text{si} \quad t \in [0, \alpha + 1/n) \\ \alpha - 1/n & \text{si} \quad t = \alpha + 1/n. \\ \text{lineal} & \text{si} \quad t \in (\alpha + 1/n, \infty). \end{cases}$$

Como es usual en el estudio de espacios métricos, se pueden definir métricas equivalentes que, en algunos casos, facilitan el estudio del espacio en cuestión. A continuación, algunas definiciones. Para cada $T > 0$ y cada par de funciones $x, y \in \mathcal{D}([0,\infty), E)$ se define la distancia $d_T(x,y)$ como el ínfimo de aquellos δ para los cuales existen $0 = t_0 < t_1 < \cdots < t_k$, con $t_k \geq T$, y $0 = s_0 < s_1 < \cdots < s_k$, con $s_k \geq T$, tales que $|t_i - s_i| \leq \delta$ para $i = 0, \ldots, k$, y

$$r\,(x(t), y(s)) \leq \delta \quad \text{si} \quad t_i \leq t \leq t_{i+1} \quad \text{y} \quad s_i \leq s \leq s_{i+1}$$

para $i = 0, \ldots, k-1$. La suma

$$\hat{d}(x,y) = \sum_{k=1}^{\infty} 2^{-k}\,(1 \wedge d_k(x,y))$$

define una métrica en el espacio $\mathcal{D}([0,\infty), E)$, la cual es equivalente a la métrica d.

Proposición 2.2.10. *La aplicación $\hat{d}$ es una métrica en el espacio $\mathcal{D}([0,\infty), E)$, la cual es equivalente a la métrica d.*

Demostración. Primero, se probará que $\hat{d}$ es una métrica. Sean $x, y \in \mathcal{D}([0,\infty), E)$, $t \in \mathbb{R}^+$ y $T > t$. Si $\hat{d}(x,y) = 0$, entonces $d_k(x,y) = 0$ para toda $k \geq 1$. Así, por definición de la distancia d_T, para toda $\epsilon > 0$, existen $\epsilon > \delta > 0$ y $k \geq 1$ para los cuales existen $0 = t_0 < t_1 < \cdots < t_k$, con $t_k \geq T$, y $0 = s_0 < s_1 < \cdots < s_k$, con $s_k \geq T$, tales que $|t_i - s_i| \leq \delta$ para $i = 0, \ldots, k$, y

$$r\,(x(t), y(s)) \leq \delta \quad \text{si} \quad t_i \leq t \leq t_{i+1} \quad \text{y} \quad s_i \leq s \leq s_{i+1}$$

para $i = 0, \ldots, k-1$,. Así, para cada τ con $t < \tau < T$, la definición de la distancia d_T justifica la existencia de una sucesión $\{s_n\}_{n\geq 1}$ tal que

$$\lim_{n\to\infty} s_n = \tau \quad \text{y} \quad \lim_{n\to\infty} y(s_n) = x(\tau).$$

Por la propiedad *càdlàg* de y en τ, $x(\tau) = y(\tau)$ o $x(\tau) = y(\tau-)$. Luego, eligiendo una sucesión $\{\tau_n\}_{n\geq 1}$ estrictamente decreciente a t, la propiedad *càdlàg* de y en t conduce a

$$\lim_{n\to\infty} y(\tau_n) = y(t) \quad \text{y} \quad \lim_{n\to\infty} y(\tau_n-) = y(t).$$

Y por este mismo argumento, $x(\tau_n)$ converge a $x(t)$ cuando $n \to \infty$. Por lo tanto, $x(t) = y(t)$. La desigualdad del triángulo se sigue directo de la definición.

Finalmente, usando la Proposición 2.2.8 (ii), la equivalencia se sigue de la siguiente afirmación: $\hat{d}(x_n, x)$ converge a 0 cuando $n \to \infty$ si, y sólo si, existe una sucesión $\{\lambda_n\}_{n\geq 1}$ de funciones continuas y estrictamente creciente de $[0,\infty)$ sobre si mismo tal que, uniformemente sobre conjuntos compactos,

$$\lim_{n\to\infty} \lambda_n(t) = t \quad \text{y} \quad \lim_{n\to\infty} r\left(x\big(\lambda_n(t)\big), x_n(t)\right) = 0.$$

■

Teorema 2.2.11. *Si E es un espacio separable, entonces $\mathcal{D}([0,\infty), E)$ es separable. Si (E, r) es un espacio métrico completo, entonces $(\mathcal{D}([0,\infty), E), d)$ es un espacio métrico completo.*

Demostración. (Separabilidad). Sea E separable y $\{\alpha_1, \alpha_2, \ldots\} \subset E$ un conjunto denso numerable. Sea Γ la colección numerable de elementos de $\mathcal{D}([0,\infty), E)$ de la forma

$$y(t) := \begin{cases} \alpha_{i_{k-1}} & \text{si} \quad t_{k-1} \leq t < t_k, \quad k = 1, \ldots, n; \\ \\ \alpha_{i_n} & \text{si} \qquad\qquad t \geq t_n, \end{cases}$$

donde $0 = t_0 < t_1 < \ldots < t_n$ son racionales, $i_1, \ldots, i_n$ son enteros positivos y $n \geq 1$. Se probará que Γ es un conjunto denso en $\mathcal{D}([0,\infty), E)$. Sea $\epsilon > 0$, $x \in \mathcal{D}([0,\infty), E)$ y $T \in \mathbb{N}$ suficientemente grande tal que $\exp\{-T\} < \epsilon/2$. Por el Lema 2.2.1, x tiene una cantidad numerable de puntos discontinuos, los cuales se denotarán por $s_1, \ldots, s_m$ para algún $m \in \mathbb{N}$ con $0 = s_0 < s_1 < \ldots < s_m < s_{m+1} = T$. Además, ya que x es continua por la derecha, para cada $j = 0, 1, \ldots, m$, x es uniformemente continua en $[s_j, s_{j+1})$, por lo que existe $0 < \delta_j < s_{j+1} - s_j$ tal que, si $s, t \in [s_j, s_{j+1})$ con $|s - t| < \delta_j$, entonces

$$r(x(s), x(t)) < \frac{\epsilon}{4}. \tag{2.2.16}$$

Sean

$$n \geq \frac{3T}{\delta_0 \wedge \cdots \wedge \delta_m} \quad \text{y} \quad t_k = \frac{kT}{n}$$

para $k = 0, \ldots, n$. Ahora, por la separabilidad de E, para cada $k = 0, \ldots, n-1$, existe i_k tal que

$$r(\alpha_{i_k}, x(t_k)) < \frac{\epsilon}{4}. \tag{2.2.17}$$

Sea $y \in \Gamma$ con valores en $\alpha_{i_1}, \ldots, \alpha_{i_{n-1}}$. Para cada $j = 1, \ldots, m$ con $ns_j \notin \mathbb{N}$, existe algún k_j tal que $s_j \in (t_{k_j}, t_{k_j+1})$. Sean

$$t'_{k_j} = \frac{s_j - \exp\{-\epsilon\} t_{k_j+1}}{1 - \exp\{-\epsilon\}}, \quad \text{y} \quad t'_{k_j+1} = \frac{s_j - \exp\{\epsilon\} t_{k_j+1}}{1 - \exp\{\epsilon\}},$$

los cuales pertencen a (t_{k_j-1}, s_j) y (t_{k_j+1}, t_{k_j+2}), respectivamente. Por la elección de n, $|k_i - k_j| \geq 3$ si $j \neq i$, por lo que t'_{k_j} es estrictamente creciente. Sea $\lambda \in \Lambda$ la función lineal a trozos estrictamente creciente con

$$\lambda(t'_{k_j}) = t'_{k_j}, \quad \lambda(t_{k_j+1}) = s_j \quad \text{y} \quad \lambda(t'_{k_j+1}) = t'_{k_j+1},$$

para aquellos $j = 1, \ldots, m$ para los cuales $ns_j \notin \mathbb{N}$, y lineal con pendiente 1 para $t > t'_{k+j}$. Esta última condición implica

$$\begin{aligned}
\gamma(\lambda) &\leq \min_j \left\{ \log \left| \frac{t'_{k_j+1} - s_j}{t'_{k_j+1} - t_{k_j+1}} \right| \wedge \log \left| \frac{s_j - t'_{k_j}}{t_{k_j+1} - t'_{k_j}} \right| \right\} \\
&\leq \min_j \left\{ \log |\exp\{\epsilon\}| \wedge \log |\exp\{-\epsilon\}| \right\}
\end{aligned}$$

siendo este último término igual a ϵ,. Si $t < T$, entonces, por (2.2.16), (2.2.17) y la definición de y,

$$r\big(x(t), y(\lambda(t)\big) \leq r\big(x(t), x(t_k)\big) + r\big(x(t_k), \alpha_{i_k}\big) < \frac{\epsilon}{2} \tag{2.2.18}$$

para alguna k con $t - t_k < \delta = \delta_1 \wedge \ldots \wedge \delta_m$. Luego, por (2.2.18), la definición de d y la elección de T,

$$d(x, y) < \epsilon \vee \left(\frac{\epsilon}{2} + \exp\{-T\} \right) \leq \epsilon.$$

Por lo tanto, Γ es un conjunto denso en $\mathcal{D}([0, \infty), E)$.

(Completud). Es suficiente verificar que toda sucesión de Cauchy tiene una subsucesión que converge. Si $\{x_n\}_{n \geq 1} \subset \mathcal{D}([0, \infty), E)$ es una sucesión de Cauchy, entonces, para toda $k \in \mathbb{N}$, existe N_k tal que

$$d(x_m, x_n) \leq 2^{-(k+1)} \exp\{-k\},$$

para toda $n, m \geq N_k$. Esto es, para $k \geq 1$, se puede elegir $\lambda_k \in \Lambda$ y $u_k > k$ (pues $\exp\{-u\} d'(x, y, \lambda, u)$ está acotada) tal que

$$\gamma\left(\lambda_k\right) \vee \sup_{t \geq 0} q\big(x_{N_k}(t \wedge u_k), x_{N_{k+1}}(\lambda_n(t) \wedge u_k)\big) \leq 2^{-k}. \tag{2.2.19}$$

Sea $\mu_{n,k} = \lambda_{k+n+1} \circ \lambda_{k+n} \circ \ldots \circ \lambda_{k+1} \circ \lambda_k$. Por otro lado, como $|u - 1| \leq \exp\{|\log u|\} - 1$ para $u \geq 0$, entonces, para todo intervalo acotado I,

$$\sup_{t \in I} |\lambda(t) - t| = \sup_{t \in I \setminus \{0\}} t \left| \frac{\lambda(t) - \lambda(0)}{t - 0} - 1 \right| \leq C\big(\exp\{\gamma\left(\lambda\right)\} - 1\big), \tag{2.2.20}$$

donde C es una constante no negativa que depende de I. Este razonamiento permite concluir, para cualquier $\mu, \lambda \in \Lambda$,

$$\sup_{t \neq \lambda(t)} \left| \frac{\mu(\lambda(t)) - \mu(t)}{\lambda(t) - t} \right| \leq \exp\{\gamma(\mu)\}.$$

Más aún, para todo intervalo acotado I,

$$\sup_{t \in I} |\mu(\lambda(t)) - \mu(t)| \leq \sup_{t \in I} |\lambda(t) - t| \exp\{\gamma(\mu)\}. \tag{2.2.21}$$

Así, si $n \geq m$, entonces, por (2.2.21), (2.2.20) y $\gamma(\lambda_k) < 2^{-k}$,

$$\begin{aligned}
\sup_{t \in I} |\mu_{n+1,k}(t) - \mu_{n,k}(t)| &\leq \sup_{t \in I} |\mu_{n-m-1,k+m+1}(t) - t| \exp\{\gamma(\mu_{m,k})\} \\
&\leq C \left(\exp\{\mu_{n-m-1,k+m+1}\} - 1\right) \exp\{2^{-k+1}\} \\
&\leq C \left(\exp\{2^{-k+1}\} - 1\right) \exp\{2^{-k+1}\}
\end{aligned}$$

para todo intervalo acotado I y alguna constante no negativa C que depende de I. Luego,

$$\lim_{m \to \infty} \sup_{t \in I} |\mu_{n+1,k}(t) - \mu_{n,k}(t)| = 0.$$

Esto significa que, para k fijo, $\{\mu_{n,k}\}_{n \geq 1}$ es una sucesión de Cauchy en todo intervalo acotado. Por lo tanto, la sucesión anterior converge uniformemente sobre todo intervalo acotado. Sea

$$\mu_k := \lim_{n \to \infty} \lambda_{k+n} \circ \ldots \circ \lambda_{k+1} \circ \lambda_k.$$

Se observa que, para cada $k \geq 1$, la función μ_k es continua y creciente, propiedades que hereda de la sucesión $\{\lambda_k\}_{k \geq 1}$. Además, dado que

$$\gamma(\mu_k) = \sup_{0 \leq s < t} \left| \log \frac{\mu_k(t) - \mu_k(s)}{t - s} \right| \leq \sup_{0 \leq s < t} \sum_{n=0}^{\infty} \left| \log \frac{\lambda_{k+n} \circ \cdots \circ \lambda_k(t) - \lambda_{k+n} \circ \cdots \circ \lambda_k(s)}{\lambda_{k+n-1} \circ \cdots \circ \lambda_k(t) - \lambda_{k+n-1} \circ \cdots \circ \lambda_k(s)} \right|$$

con $\lambda_{k-1} \circ \lambda_k$ como la función identidad, el Lema 2.2.3 (ii) y (2.2.19) justifican

$$\gamma(\mu_k) \leq \lim_{n \to \infty} \sum_{i=k}^{n} \gamma(\lambda_i) = \sum_{i=k}^{\infty} \gamma(\lambda_i) \leq 2^{-k+1},$$

y así, μ_k es Lipschitz continua. En consecuencia, para cada $k \in \mathbb{N}$, $\mu_k \in \Lambda$. Más aún, por la desigualdad anterior se deduce

$$\lim_{k \to \infty} \gamma(\mu_k) = 0. \tag{2.2.22}$$

Por otra parte, para toda $k \geq 1$, la continuidad de λ_k y (2.2.19) implican

$$\begin{aligned}
&\sup_{t \geq 0} q\left(x_{N_k}\left(\mu_k^{-1}(t) \wedge u_k\right), x_{N_{k+1}}\left(\mu_{k+1}^{-1}(t) \wedge u_k\right)\right) \\
&\quad = \sup_{t \geq 0} q\left(x_{N_k}\left(\mu_k^{-1}(t) \wedge u_k\right), x_{N_{k+1}}\left(\lambda_k\left(\mu_k^{-1}(t)\right) \wedge u_k\right)\right) \\
&\quad = \sup_{t \geq 0} q\left(x_{N_k}\left(t \wedge u_k\right), x_{N_{k+1}}\left(\lambda_k(t) \wedge u_k\right)\right) \\
&\quad \leq 2^{-k}.
\end{aligned}$$

Esta desigualdad, junto con la completud de E, asegura la convergencia uniforme sobre todo intervalo acotado de $x_{N_k} \circ \mu_k^{-1}$ a una función $x : [0, \infty) \to E$. Y ya que cada $x_{N_k} \circ \mu_k^{-1}$ pertenece a $\mathcal{D}([0, \infty), E)$, x también debe pertenecer[8] a $\mathcal{D}([0, \infty), E)$. Como resultado se obtiene

$$\lim_{k\to\infty} \sup_{0\leq t\leq T} r\left(x_{N_k}\left(\mu_k^{-1}(t)\right), y(t)\right) = 0$$

para toda T real positivo. Se sigue de (2.2.22) y la Proposición 2.2.8 (ii) que

$$\lim_{k\to\infty} d(x_{N_k}, x) = 0.$$

■

Teorema 2.2.12. *Para cada $t \geq 0$ se define $\pi_t : \mathcal{D}([0, \infty), E) \to E$ por $\pi_t(x) = x(t)$. Sean $\mathcal{F}_E$ la $\sigma-$álgebra de Borel de $\mathcal{D}([t, \infty), E)$ y*

$$\mathcal{F}'_E := \sigma\left\langle \pi_t \;\middle|\; 0 \leq t < \infty \right\rangle = \sigma\left\langle \pi_s \;\middle|\; s \in D \right\rangle, \tag{2.2.23}$$

donde D es cualquier conjunto denso en $[0, \infty)$. Si E es un espacio separable, entonces $\mathcal{F}_E = \mathcal{F}'_E$.

Demostración. Primero se verificará (2.2.23). Para ello, es suficiente probar

$$\mathcal{F}'_E \subset \sigma\left\langle \pi_s \,|\, s \in D \right\rangle.$$

Para cada $t \geq 0$, existe $\{t_n\}_{n\geq 1} \subset D \cap [0, \infty)$ tal que $\lim_{n\to\infty} t_n = t$. Luego, por la continuidad por la derecha de x,

$$\lim_{n\to\infty} \pi_{t_n}(x) = \lim_{n\to\infty} x(t_n) = x(t) = \pi_t(x).$$

Y como π_{t_n} es $\sigma\left\langle \pi_s \;\middle|\; s \in D \right\rangle -$medible, π_t también lo es. (2.2.23) se sigue.

Ahora, se probará $\mathcal{F}'_E \subset \mathcal{F}_E$. Sean $x, y \in \mathcal{D}([0, \infty), E)$. Dado $\epsilon > 0$, $t \in [0, \infty)$ y f una función real-valuada continuo y acotada en E, se define en $\mathcal{D}([0, \infty), E)$ la siguiente aplicación

$$f_t^\epsilon(x) := \frac{1}{\epsilon} \int_t^{t+\epsilon} f(\pi_s(x))ds,$$

Ahora, sean $x_n, x \in \mathcal{D}([t, \infty), E)$, con $n \geq 1$, tales que x_n converge a x cuando $n \to \infty$. Entonces, por el Lema 2.2.4, existe una sucesión $\{\lambda_n\}_{n\geq 1} \subset \Lambda$ tal que $\lim\limits_{n\to\infty} \gamma(\lambda_n) = 0$ y

$$\lim_{n\to\infty} m\left(u \in [0, T] \;\middle|\; d'(x_n, y_n, \lambda_n, u) \geq \epsilon\right) = 0$$

para toda $\epsilon > 0$ y toda $T > 0$, donde m denota la medida de Lebesgue. Y ya que $x_n(\lambda_n(s))$ y $\lambda_n(s)$ converge uniformemente a $x(s)$ y a s sobre intervalos acotados, respectivamente, y $\lambda_n(t)$ converge casi en todas partes uniformemente a 1 sobre intervalos acotados, entonces, por el teorema de convergencia dominada y la continuidad de f,

$$\lim_{n\to\infty} f_t^\epsilon(x_n) = \lim_{n\to\infty} \frac{1}{\epsilon} \int f\left(x_n(\lambda_n(s))\right) \mathbf{1}_{\{\lambda_n(t)\leq s\leq\lambda_n(t+\epsilon)\}} \lambda_n(s)' ds = f_t^\epsilon(x),$$

[8]Límite uniforme preserve la propiedad *càdlàg*.

donde la primera igualdad es consecuencia de un cambio de variable. Este hecho implica que $f_t^\epsilon(x)$ es una función Borel medible. Más aún, por L'Hospital y la continuidad por la derecha de x, se tiene que

$$\lim_{\epsilon\to\infty} f_t^\epsilon(x) = f(\pi_t(x)),$$

lo que permite deducir que $f \circ \pi_t$ es una función Borel medible para toda f función real-valuada continuo y acotada en E, y así, por el Teorema de clases monónotas, para toda función medible y acotada. En particular, dado que $\mathbf{1}_A(\pi_t(x))$ es Borel medible para toda $A \in \mathcal{B}(E)$,

$$\pi_t^{-1}(A) := \big\{\omega \in \mathcal{D}([0,\infty), E) \,|\, \pi_t(\omega) \in A\big\} \in \mathcal{F}_E,$$

para toda $A \in B(E)$, lo que prueba $\mathcal{F}'_E \subset \mathcal{F}_E$.

Recíprocamente, para cada $n \geq 1$, $0 = t_0 < t_1 < \ldots < t_n < \infty$ y $\alpha_0, \ldots, \alpha_n \in E$, se define la aplicación $\eta : E^{n+1} \to \mathcal{D}([0,\infty), E)$ dada por

$$\eta\left(\alpha_0, \ldots, \alpha_n\right)(t) = \sum_{k=0}^{n-1} \alpha_k \mathbf{1}_{[t_k, t_{k+1})}(t) + \mathbf{1}_{[t_n \infty)}(t).$$

Se recuerda que r es una métrica en E, con E un espacio separable. Entonces, para $\left(\alpha_0, \ldots, \alpha_n\right), \left(\alpha'_0, \ldots, \alpha'_n\right) \in E^{n+1}$ y $u \in [0,\infty)$, se cumple

$$\sup_{t\geq 0} q\left(\eta\left(\alpha_0, \ldots, \alpha_n\right)(t \wedge u), \eta\left(\alpha'_0, \ldots, \alpha'_n\right)(\lambda(t) \wedge u)\right) \leq \max_{0\leq i\leq n} r(\alpha_i, \alpha'_i).$$

Esto implica

$$d\left(\eta\left(\alpha_0, \ldots, \alpha_n\right), \eta\left(\alpha'_0, \ldots, \alpha'_n\right)\right) \leq \max_{0\leq i\leq n} r(\alpha_i, \alpha'_i).$$

Y debido a que $\max_{0\leq i\leq n} r(\alpha_i, \alpha'_i)$ es una métrica en E^{n+1}, la desigualdad anterior conduce a la continuidad de η. Más aún, ya que π_t es por definición $\mathcal{F}'_E-$medible y E es separable, para cada $z \in \mathcal{D}([0,\infty), E)$ y $n \geq 1$, la aplicación $\kappa_{z,(t_0,\ldots,t_n)} : \mathcal{D}([0,\infty), E) \to \mathbb{R}$ dada por

$$x \mapsto \left(\pi_{t_1}, \ldots, \pi_{t_n}\right)(x) \mapsto \eta \circ \left(\pi_{t_1}, \ldots, \pi_{t_n}\right)(x) \mapsto d\left(z, \eta\big(\pi_{t_0}(x), \ldots, \pi_{t_n}(x)\big)\right)$$

es $\mathcal{F}'_E-$medible[9]. Finalmente, para cada $m \in \mathbb{N}$, sea η_m definida como η con $n = m^2$ y $t_k = k/m$, $k = 0, \ldots, m^2$. Entonces, por un argumento similar al desarrollado en la parte de separabilidad del Teorema 2.2.11, para toda $x \in \mathcal{D}([0,\infty), E)$, se tiene

$$\lim_{m\to\infty} d(x, \eta_m(\pi_{t_0}(x), \ldots, \pi_{t_n}(x))) = 0.$$

En consecuencia,

$$\lim_{m\to\infty} |d(x, \eta_m(\pi_{t_0}(x), \ldots, \pi_{t_n}(x))) - d(z,x)| \leq \lim_{m\to\infty} d(x, \eta_m(\pi_{t_0}(x), \ldots, \pi_{t_n}(x))) = 0$$

[9]Más precisamente, sean $z \in \mathcal{D}([0,\infty), E)$, A un subonjunto denso y numerable en E^{n+1} y $\epsilon > 0$ dados. Entonces, por la continuidad de η,

$$\Gamma = \big\{a \in E^{n+1} \,\big|\, d(z, \eta(a)) < \epsilon\big\} = \bigcup_{\left\{a\in A \,\middle|\, d(z,\eta(a))<\epsilon\right\}} \bigcap_{n\geq 1} \left\{b \in E^{n+1} \,\middle|\, r'(a,b) \leq 1/n\right\}$$

es un subconjunto medible en E^{n+1} con respecto a la sigma álgebra de Borel producto, donde r' denota un producto en E^{n+1}.

para toda $x \in \mathcal{D}([0,\infty),E)$. Por lo tanto,

$$\lim_{m\to\infty} d\left(z, \eta\big(\pi_{t_0}(x),\ldots,\pi_{t_{m^2}}(x)\big)\right) = d(z,x).$$

Así, la aplicación $x \mapsto d(z,x)$ es $\mathcal{F}'_E-$medible para toda $z \in \mathcal{D}([0,\infty),E)$. En particular, toda bola abierta, con centro z y radio $\epsilon > 0$,

$$B(z,\epsilon) := \{x \in \mathcal{D}([0,\infty),E) \,|\, d(z,x) < \epsilon\}$$

pertenece a $\mathcal{F}'_E$, y ya que $\mathcal{D}([0,\infty),E)$ es separable (gracias a la separabilidad de E y al Teorema 2.2.11), $\mathcal{F}'_E$ contiene a todos los conjuntos abiertos en $\mathcal{D}([0,\infty),E)$, y en consecuencia, contiene a $\mathcal{F}_E$.

■

2.2.1. Compacidad en $\mathcal{D}([0,\infty),E)$

Como es costumbre, (E,r) es un espacio métrico. Con el fin de aplicar el Teorema 2.1.22 $\mathcal{P}(\mathcal{D}([0,\infty),E))$, se necesita dar una caracterización de los subconjuntos compactos de $\mathcal{D}([0,\infty),E)$. Con esto en mente, se dará condiciones para los cuales una colección de funciones escalonadas es compacta. Para una función escalonada $x \in \mathcal{D}([0,\infty),E)$ se define $s_0(x) = 0$ y

$$s_k(x) = \inf\{t > s_{k-1}(x) \mid x(t) \neq x(t-)\},$$

para $k = 1,2,\ldots$ si $s_{k-1}(x) < \infty$. Si $s_{k-1}(x) = \infty$, entonces $s_k(x) = \infty$.

Lema 2.2.13. *Sean $\Gamma \subset E$ un subconjunto compacto y $\delta > 0$. Se define $A(\Gamma,\delta)$ el conjunto de funciones escalonadas $x \in \mathcal{D}([0,\infty),E)$ tales que $x(t) \in \Gamma$ para toda $t \geq 0$, y $s_k(x) - s_{k-1}(x) > \delta$ para cada $k \geq 1$ para los cuales $s_{k-1}(x) < \infty$. Entonces la cerradura de $A(\Gamma,\delta)$ es compacta.*

Demostración. Es suficiente demostrar que toda sucesión en $A(\Gamma,\delta)$ tiene una subsucesión convergente. Sea $\{x_n\}_{n\geq 1} \subset A(\Gamma,\delta)$. Entonces, existe una subsucesión $\{x'_n\}_{n\geq 1}$ tal que, o bien, $s_1(x'_n) < \infty$ para toda $n \in \mathbb{N}$, o $s_1(x'_n) = \infty$ para toda $n \in \mathbb{N}$. En el primer caso, existe una subsucesión[10] de $\{x'_n\}_{n\geq 1}$, $\{x''_n\}_{n\geq 1}$, tal que, para alguna $t_1 \in [\delta,\infty]$,

$$\lim_{n\to\infty} s_1(x''_n) = t_1.$$

Si $t_1 < \infty$, entonces, para n suficientemente grande, se puede establecer la condición

$$\left|\log \frac{t_1}{s_1(x''_n)}\right| < \frac{1}{n}.$$

Ahora, ya que Γ es compacto, la sucesión $\{x''_n\}_{n\geq 1}$ tiene una subsucesión $\{x^1_n\}_{n\geq 1}$ tal que

$$\lim_{n\to\infty} x^1_n\left(s_1(x^1_n)\right) = \alpha_1$$

para alguna α_1. Prosiguiendo con este razonamiento, se puede construir una sucesión de subsucesiones $\{x^1_n\}_{n\geq 1} \subset \{x^2_n\}_{n\geq 1} \subset \cdots$ tales que, para $k = 1,2,\ldots,$

[10]Toda sucesión tiene una subsucesión convergente o divergente.

- si $s_k(x_n^k) < \infty$ para toda $n \in \mathbb{N}$, existe alguna $t_k \in [k\delta, \infty]$ para el cual se cumple $\lim_{n\to\infty} s_k(x_n^k) = t_k$, y si $t_k < \infty$,

$$\left|\log \frac{t_k - t_{k-1}}{s_k(x_n^k) - s_{k-1}(x_n^k)}\right| < \frac{1}{n} \tag{2.2.24}$$

para n suficientemente grande, y, para alguna α_k

$$\lim_{n\to\infty} x_n^k(s_k(x_n^k)) = \alpha_k \tag{2.2.25}$$

- o bien, $s_k(x_n^k) = \infty$ para toda $n \in \mathbb{N}$.

Se considera el primer caso. Sea $\{y_m\}_{m\geq 1}$ la subsucesión de $\{x_n\}_{n\geq 1}$ dada por $y_m = x_m^m$, y sea $y \in \mathcal{D}([0,\infty), E)$ definida por $y(t) = \alpha_k$ para $t_k \leq t < t_{k+1}$, $k = 0, 1, \ldots$, y con $t_0 = 0$. Como $s_k(y_m) - s_{k-1}(y_m) > \delta$ para cada k y cada m para el cual se cumple $s_{k-1}(y_m) < \infty$, se define, para cada m, la aplicación $\lambda_m \in \Lambda$ como una función lineal a trozos con

$$\lambda_m(s_{k-1}(y_m)) = t_{k-1} \quad \text{y} \quad \lambda_m(s_k(y_m)) = t_k,$$

si $t_k < \infty$, y lineal con pendiente 1 para $t > s_k(y_m)$. Entonces, por (2.2.24),

$$\gamma(\lambda_m) = \sup_{0\leq s<t}\left|\log \frac{\lambda_m(t) - \lambda_m(s)}{t-s}\right| \leq \sup_k \left|\log \frac{t_k - t_{k-1}}{s_k(y_m) - s_{k-1}(y_m)}\right| \leq \frac{1}{m} \tag{2.2.26}$$

Por otro lado, para cada $t \in \mathbb{R}^+$, existe una $k \in \{0, 1, \ldots\}$ tal que $s_k(y_m) \leq t < s_{k+1}(y_m)$ y $\lambda_m(s_k(y_m)) \leq \lambda_m(t) \leq \lambda_m(s_{k+1}(y_m))$. Entonces, por las defenciones de y_m y y,

$$r\big(y_m(t), y(\lambda_m(t))\big) = r\big(y_m(s_k(y_m)), \alpha_k\big),$$

la cual, por (2.2.25), converge a cero cuando $m \to \infty$. Luego, por la Proposición 2.2.8 (ii), y_m converge, bajo la métrica d, a y.
Ahora, se considera el segundo caso, es decir, $s_k(x_n^k) = \infty$ para toda $n \in \mathbb{N}$ y toda k. Por definición de s_k, para toda $n \geq 1$, la subsucesión $\{x_n^k\}_{k\geq 1}$ está conformada por funciones constantes, constantes que pertenecen a Γ. Luego, usando la compacidad de Γ, existe una subsecesión de $\{x_n^k\}_{k\geq 1}$ que converge.

■

Las condiciones de compacidad se expresan en términos del siguiente módulo de continuidad. Para cada $x \in \mathcal{D}([0,\infty), E)$, $\delta > 0$ y $T > 0$, se define el **módulo de continuidad**

$$w'(x, \delta, T) := \inf_{\{t_i\}} \max_i \sup_{s,t\in[t_{i-1},t_i)} r\big(x(s), x(t)\big),$$

donde el ínfimo se toma sobre todas las paritciones $\{t_i\}_{i\geq 1}$ tales que

$$t_0 < t_1 < \cdots < t_{n-1} < T < t_n \tag{2.2.27}$$

con $\min_{1\leq i\leq n}(t_i - t_{i-1}) > \delta$ y $n \geq 1$. Se observa que $w'(x, \delta, T)$ es no decreciente en δ y en T, y que

$$w'(x, \delta, T) \leq w'(y, \delta, T) + 2 \sup_{0\leq s<T+\delta} r(x(s), y(s)). \tag{2.2.28}$$

Sean $0 < \delta_1 < \delta_2$. Se denota por A_{δ_i}, $i = 1, 2$, al conjunto de particiones $\{t_i\}$ de la forma (2.2.27) con $\delta = \delta_i$, $i = 1, 2$. Entonces, la monotonía de $w'(x, \delta, T)$ en δ se sigue De $A_{\delta_2} \subset A_{\delta_1}$. Un estudio similar se lleva a cabo para probar la monotonía de $w'(x, \delta, T)$ en T.

Lema 2.2.14. **(i)** *Para cada $x \in \mathcal{D}([0,\infty), E)$ y $T > 0$, $w'(x,\delta,T)$ es continua por la derecha en δ y*
$$\lim_{\delta \to 0} w'(x,\delta,T) = 0.$$

(ii) *Si $\{x_n\}_{n\geq 1} \subset \mathcal{D}([0,\infty), E)$ y $\lim_{n\to\infty} d(x_n,x) = 0$, entonces*
$$\limsup_{n\to\infty} w'(x_n,\delta,T) \leq w'(x,\delta,T+\epsilon)$$
para toda $\delta, T, \epsilon > 0$.

(iii) *Para cada $\delta > 0$ y $T > 0$, $w'(x,\delta,T)$ es Borel medible en x.*

Demostración. (i) Sean $x \in \mathcal{D}([0,\infty), E)$ y $T > 0$ dados. Primero se demostrará la continuidad por la derecha. Sea $\delta > 0$ dado y $\delta' = \frac{1}{2}\left(\delta + \min_{1\leq i\leq n}(t_i - t_{i-1})\right) > \delta$. Entonces, para toda $\eta \in (\delta,\delta')$, la monotonía de $w'(x,\delta,T)$ en δ implica
$$w'(x,\delta,T) \leq w'(x,\eta,T).$$

Ahora se demostrará la segunda afirmación. Para $N \in \mathbb{N}$ se define $\tau_0^N := 0$, y para toda $k \geq 1$,
$$\tau_k^N := \inf\left\{t > \tau_{k-1} \mid r\left(x(t), x(\tau_{k-1}^N)\right) > \frac{1}{N}\right\}$$
si $\tau_{k-1}^N < \infty$, $\tau_k^N = \infty$ si $\tau_{k-1}^N = \infty$. Se recuerda, por la prueba dada en la Proposición 2.2.8, que $\{\tau_k^N\}_{k\geq 1}$ es estrictamente creciente. Entonces, para
$$0 < \delta < \min\left(\tau_{k+1}^N - \tau_k^N \,\middle|\, \tau_k^N < T\right),$$
$$w'(x.\delta,T) \leq \max_i \sup_{s,t\in[\tau_i^N - \tau_{i-1}^N)} \Big(r\left(x(s), x(\tau_i^N)\right) + r\left(x(\tau_i^N), x(t)\right)\Big),$$
la cual, por definición de τ_k^N, es menor o igual que $2/N$. Luego, tomando el límite cuando $N \to \infty$ en ambos lados de la desigualdad anterior, se concluye
$$\lim_{N\to\infty} w'(x,\delta,T) \leq \lim_{N\to\infty} \frac{2}{N} = 0,$$
lo que implica
$$\lim_{\delta\to 0} w'(x,\delta,T) = 0.$$

(ii) Sean $x, x_n \in \mathcal{D}([0,\infty), E)$, $n \geq 1$. También, sean δ y T reales positivos dados. Se asume que $\lim_{n\to\infty} d(x_n,x) = 0$. Entonces, por la Proposición 2.2.8 (ii), existe una sucesión $\{\lambda_n\}_{n\geq 1} \subset \Lambda$ tal que $\lim_{n\to\infty} \gamma(\lambda_n) = 0$ y
$$\lim_{n\to\infty} \sup_{0\leq t\leq T} r\big(x_n(t), x(\lambda_n(t))\big) = 0$$
para toda $T > 0$. En particular, el límite anterior se cumple considerando $T + \delta$ en vez de T, es decir,
$$\lim_{n\to\infty} \sup_{0\leq t\leq T+\delta} r\big(x_n(t), x(\lambda_n(t))\big) = 0 \tag{2.2.29}$$

Ahora, para cada $n \in \mathbb{N}$, se definen $y_n(t) := x(\lambda_n(t))$ para toda $t \in \mathbb{R}^+$ y

$$\delta_n := \sup_{0 \le t \le T} \big(\lambda_n(t+\delta) - \lambda_n(t)\big).$$

Gracias a la Proposición 2.2.3 (iii), se observa que

$$\delta = \lim_{n\to\infty} \delta_n. \tag{2.2.30}$$

Entonces, para todo $\epsilon > 0$, (2.2.28) y (2.2.29) permiten deducir

$$\begin{aligned}\limsup_{n\to\infty} w'(x_n,\delta,T) &\le \limsup_{n\to\infty} w'(y_n,\delta,T) + 2\limsup_{n\to\infty}\sup_{0\le t\le T+\delta} r\big(x_n(s), x(\lambda_n(t))\big)\\ &\le \limsup_{n\to\infty} w'(y_n,\delta,T).\end{aligned}$$

Más aun, sustituyendo s,t por $\lambda_n(s), \lambda_n(t)$ en la definición de w' y usando (2.2.30), la definición de y_n y la monotonía de $w'(x,\delta,T)$ en δ conducen a

$$\limsup_{n\to\infty} w'(x_n,\delta,T) \le \limsup_{n\to\infty} w'(y_n,\delta,T) \le \limsup_{n\to\infty} w'(x,\delta_n+1/n,\lambda_n(T))$$

Ahora bien, ya que $\lambda_n(T)$ converge a T en intervalos acotados, cuando $n \to \infty$, la monotonía de w' implica

$$\begin{aligned}\limsup_{n\to\infty} w'(x_n,\delta,T) &\le \limsup_{n\to\infty} w'(x,\delta_n+1/n,\lambda_n(T))\\ &\le \limsup_{n\to\infty} w'\left(x, (\delta_n+1/n)\vee(\delta+1/n), T+\epsilon\right).\end{aligned}$$

Luego, por la continuidad por la derecha de $w'(x,\delta,T)$ en δ,

$$\begin{aligned}\limsup_{n\to\infty} w'(x_n,\delta,T) &\le \limsup_{n\to\infty} w'\left(x, (\delta_n+1/n)\vee(\delta+1/n), T+\epsilon\right)\\ &\le \lim_{n\to\infty} w'\left(x, (\delta_n+1/n)\vee(\delta+1/n), T+\epsilon\right)\\ &\le w'(x,\delta,T+\epsilon),\end{aligned}$$

Por lo tanto,

$$\limsup_{n\to\infty} w'(x_n,\delta,T) \le w'(x,\delta,T+\epsilon).$$

(iii) Sean δ y T reales positivos dados. Se define $w'(x,\delta,T+) := \lim_{\epsilon\downarrow 0} w'(x,\delta,T+\epsilon)$, la cual existe por la monotonía de w' en T (de hecho la sucesión en cuestión es decreciente y acotado por abajo). Sean $x, x_n \in \mathcal{D}([0,\infty),E)$ para $n \ge 1$ tales que

$$\lim_{n\to\infty} d(x_n,x) = 0.$$

Entonces, por al apartado (ii) de este Lema,

$$\begin{aligned}\limsup_{n\to\infty} w'(x_n,\delta,T+) &= \limsup_{n\to\infty}\lim_{\epsilon\downarrow 0} w'(x_n,\delta,T+\epsilon)\\ &\le \lim_{\epsilon\downarrow 0} w'(x,\delta,T+2\epsilon)\\ &= w'(x,\delta,T+).\end{aligned}$$

Así, $w'(x, \delta, T+)$ es semicontinua superiormente (upper semicontinuous)[11], y por un resultado clásico de análisis, es Borel medible respecto a x. El resultado se sigue de la siguiente observación:

$$w'(x, \delta, T) = \lim_{\epsilon \to 0} w'(x, \delta, (T - \epsilon)+)$$

para toda $x \in \mathcal{D}([0, \infty), E)$. ■

Teorema 2.2.15. *Sea (E, r) completo. La cerradura de $A \in \mathcal{D}([0, \infty), E)$ es compacta si, y sólo si, se cumple*

(i) *para cada racional $t \geq 0$, existe un conjunto compacto, $\Gamma_t \subset E$, tal que $x(t) \in \Gamma_t$ para toda $x \in A$;*

(ii) *para cada $T > 0$,*

$$\lim_{\delta \to 0} \sup_{x \in A} w'(x, \delta, T) = 0.$$

Demostración. Sea $A \in \mathcal{D}([0, \infty), E)$ tal que cumpla las condiciones (i) y (ii) del teorema. Se probará que A es totalmente acotado y $\overline{A}$ es completo para concluir que A es relativamente compacto. Es claro que la completitud de $\mathcal{D}([0, \infty), E)$ implica la completitud de $\overline{A}$. Así, para este apartado, sólo falta verificar que A es totalmente acotado.
Sean $x \in \mathcal{D}([0, \infty), E)$ y $\ell \geq 1$. Por la condición (ii) del teorema, se puede elegir $\delta_\ell \in (0, 1)$ tal que

$$\sup_{x \in A} w'(x, \delta_l, l) < \frac{1}{\ell} \tag{2.2.31}$$

y $m_\ell \geq 2$ tal que $1/m_\ell < \delta_\ell$. Sea

$$\Gamma^{(\ell)} = \bigcup_{i=0}^{(\ell+1)m_\ell} \Gamma_{\frac{i}{m_\ell}}.$$

Es claro que $x(0), x(1/m_\ell), \ldots, x(\ell+1) \in \Gamma^{(\ell)}$. Usando la notación establecida en el Lema 2.2.13, sea $A_\ell = A(\Gamma^{(\ell)}, \delta_l)$. Por otro lado, por (2.2.31) y la definición de w', existe una partición

$$0 = t_0 < t_1 < \cdots < t_{n-1} < \ell \leq t_n < \ell + 1 < t_{n+1} = \infty. \tag{2.2.32}$$

con $\min_{1 \leq i \leq n}(t_i - t_{i-1}) > \delta_\ell$ tal que

$$\max_{1 \leq i \leq n} \sup_{s, t \in [t_{i-1}, t_i)} r\big(x(s), x(t)\big) \leq \frac{2}{\ell}.$$

Ahora, se define $x' \in A_\ell$ dado por

$$x'(t) = x\left(\left(\lfloor m_\ell t_i \rfloor + 1\right)/m_\ell\right)$$

---

[11] Se recuerda que una función real valuada f es semicontinua superiomente, si

$$\limsup_{x \to x_0} f(x) \leq f(x_0).$$

para $t_i \leq t < t_{i+1}$, $i = 0, 1, \ldots, n$. Se observa que, debido a $m_\ell^{-1} < \delta_\ell$,

$$t_i \leq \left(\lfloor m_\ell t_i \rfloor + 1\right)/m_\ell \leq t_i + 1/m_\ell < t_{i+1}. \tag{2.2.33}$$

Entonces, tomando en cuenta la elección de la partición dada en (2.2.32) y (2.2.33), se obtiene que

$$\sup_{0 \leq t < \ell} r\big(x'(t), x(t)\big) \leq \frac{2}{\ell}.$$

En consecuencia,

$$\begin{aligned} d(x', x) &\leq \int_0^\infty \exp\{-u\}\Big(\sup_{t\geq 0} r\big(x'(t\wedge u), x(t\wedge u)\big) \wedge 1\Big) du \\ &= \int_0^\ell \exp\{-u\}\left(\sup_{0\leq t\leq u} r\big(x'(t), x(t)\big) \wedge 1\right) du \\ &\quad + \int_\ell^\infty \exp\{-u\}\left(\sup_{0\leq t\leq u} r\big(x'(t), x(t)\big) \wedge 1\right) du \\ &\leq \int_0^\infty \exp\{-u\}\left(\sup_{0\leq t\leq \ell} r\big(x'(t), x(t)\big) \wedge 1\right) du \\ &\quad + \int_\ell^\infty \exp\{-u\} du \\ &\leq 2/\ell + \exp\{-\ell\} < 3/\ell. \end{aligned}$$

Así, si $A_\ell^{3/\ell}$ denota la bola de radio $3/\ell$ y centro x', $A \subset A_\ell^{3/\ell}$. Luego, por la arbitrariedad de ℓ,

$$A \subset \bigcap_{\ell \geq 1} A_\ell^{3/\ell}. \tag{2.2.34}$$

Y ya que la compacidad de cada $\overline{A}_\ell$ implica que cada A_ℓ sea totalmente acotado(gracias al Lema 2.2.13), (2.2.34) permite concluir que A es totalmente acotado[12]. En efecto, sean $x \in \mathcal{D}([0,\infty), E)$ y $\epsilon > 0$. Entonces, por (2.2.34), se puede elegir ℓ para el cual satsiface $d(x, x^\ell) < \epsilon$ para algún $x^\ell \in A_\ell$. Como cada A_ℓ es totalmente acotado, existe una $\epsilon-$vecindad finita, $\{x_k^\ell\}$, con la propiedad de que para cada $x^\ell \in A_\ell$ se cumple $d(x^\ell, x_k^\ell) < \epsilon$. Así,

$$d(x, x_k^\ell) \leq d(x, x^\ell) + d(x^\ell, x_k^\ell) < 2\epsilon.$$

Por lo tanto, $\{x_k^\ell\}$ es una $\epsilon-$vecindad finito para A, lo que prueba que A es totalmente acotado.

Recíprocamente, sea A un conjunto relativamente compacto. Para cada racional no negativo t, se define $\Gamma_t \subset E$ por $\Gamma_t := \overline{A}_t$, donde $A_t = \big\{x(t) \,\big|\, x \in A\big\}$. Para probar que Γ_t es compacto, es suficiente demostrar que toda sucesión en A_t tiene una subsucesión convergente. Para cada $t \in \mathbb{R}^+$, sea $\{x_n(t)\}_{n\geq 1}$ una sucesión en A_t. Como A es relativamente compacto, restringiendo a una subsucesión si es necesario, se puede asumir que

$$\lim_{n\to\infty} d(x_n, x) = 0$$

[12]Sea (X, ρ) un espacio métrico. Se recuerda que un $\epsilon-$vecindad para $B \subset X$, es un conjunto de puntos $\{x_k\}$ con la propiedad que para cada x en B existe un x_k tal que $\rho(x, x_k) < \epsilon$. También, se recuerda que $B \subset X$ es totalmente acotado si para cada real positivo ϵ, B tiene un $\epsilon-$vecindad finita.

para algún $x \in \mathcal{D}([0,\infty),E)$. Entonces, por la Proposición 2.2.8 (ii) y el Lema 2.2.3, existe una sucesión $\{\lambda_n\}_{n\geq 1} \subset \Lambda$ tal que $\lim_{n\to\infty} \gamma(\lambda_n) = 0$,

$$\lim_{n\to\infty} \sup_{t\in[0,T]} |\lambda_n(t) - t| = 0 \tag{2.2.35}$$

y

$$\lim_{n\to\infty} \sup_{0\leq t\leq T} r\big(x_n(t), x(\lambda_n(t))\big) = 0, \tag{2.2.36}$$

para toda $T > 0$. Por otro lado, existe una subsucesión $\{\lambda_{n_r}\}_{r\geq 1}$ tal que $\lambda_{n_r}(t) \geq t$ para toda $r \geq 1$, o bien, $\lambda_{n_r}(t) < t$ para toda $r \geq 1$. En el primer caso, se tiene la siguiente desigualdad

$$r\big(x_{n_r}(t), x(t)\big) \leq r\big(x_{n_r}(t), x(\lambda_{n_r}(t))\big) + r\big(x(\lambda_{n_r}(t)), x(t)\big),$$

la cual, gracias a (2.2.36), (2.2.35) y la continuidad por la derecha de x, cumple

$$\lim_{n\to\infty} r\big(x_{n_r}(t), x(t)\big) \lim_{n\to\infty} \leq r\big(x_{n_r}(t), x(\lambda_{n_r}(t))\big) + \lim_{n\to\infty} r\big(x(\lambda_{n_r}(t)), x(t)\big) = 0.$$

En el segundo caso, se tiene

$$r\big(x_{n_r}(t), x(t-)\big) \leq r\big(x_{n_r}(t), x(\lambda_{n_r}(t))\big) + r\big(x(\lambda_{n_r}(t)), x(t-)\big),$$

la cual, por el mismo argumento del caso anterior, converge también a 0. Así, para ambos casos, se concluye que $\{x_n(t)\}_{n\geq 1}$ tiene una subsucesión convergente, por lo que la condición (i) del teorema se sostiene.

Ahora, se verificará la condición (ii) del teorema. Se procederá por reducción al absurdo. Para ello, se asume la existencia de $\eta, T > 0$ y una sucesión $\{x_n\}_{n\geq 1}$ en A tal que

$$w'(x_n, 1/n, T) \geq \eta, \quad \forall n \in \mathbb{N}.$$

Ya que A es relativamente compacto, se puede suponer que

$$\lim_{n\to\infty} d(x_n, x) = 0$$

para algún $x \in \mathcal{D}([0,\infty),E)$. Entonces, por el Lema 2.2.14 (ii),

$$\eta \leq \limsup_{n\to\infty} w'(x_n, \delta, T) \leq w'(x, \delta, T+1)$$

para toda $\delta > 0$. Luego, haciendo $\delta \to 0$, el Lema 2.2.14 (i) permite concluir

$$\eta \leq \lim_{\delta\to 0} w'(x, \delta, T+1) = 0,$$

lo que contradice la elección de η, que es positivo. Por lo tanto, la condición (ii) dada en el planteamiento de este teorema se sostiene. ■

2.2.2. Convergencia de procesos con trayectorias en $\mathcal{D}([0,\infty),E)$

Con lo desarrollado hasta ahora, se puede enunciar y probar el resultado principal de este capítulo. Este establece que una sucesión de procesos estocásticos converge débilmente si, y sólo si, la colección de medidas de probabilidad inducidas es relativamente compacto y las distribuciones finito dimensionales convergen, la cual es un método útil para probar convergencia de procesos estocásticos.
Se recuerda que un proceso estocástico X es una familia de elementos aleatorios definidas en un espacio de probabilidad $(\Omega,\mathcal{F},\mathbf{P})$ y con valores en un espacio métrico, en nuestro caso, con valores en E. Se recuerda que $\mathcal{F}_E$ denota la $\sigma-$álgebra de Borel de $\mathcal{D}([0,\infty),E)$. Entonces, la **distribución del proceso** X está dada por

$$\mathbf{P}X^{-1}(B) := \mathbf{P}\left(X\in B\right),\ B\in\mathcal{F}_E.$$

Así, el estudio de convergencia débil de procesos estocásticos se reduce a estudiar convergencia débil de las distribuciones de los procesos en cuestión.

Definición 2.2.16. *Sea $\{X^\alpha,\alpha\in I\}$(para algún conjunto de índice I) una familia de procesos estocásticos con trayectorias en $\mathcal{D}([0,\infty),E)$ y $\{\mathbf{P}_\alpha\}_{\alpha\in I}\subset\mathcal{P}\left(\mathcal{D}([0,\infty),E)\right)$ la familia de sus distribuciones asociada, es decir,*

$$\mathbf{P}_\alpha\left(B\right) := \mathbf{P}\left(X^\alpha\in B\right),\ B\in\mathcal{F}_E,$$

donde $\mathbf{P}$ es la medida de probabilidad del espacio de probabilidad donde se encuentra definido el proceso X_α. Se dice que la familia $\{X^\alpha,\alpha\in I\}$ es relativamente compacta si $\{\mathbf{P}_\alpha\}_{\alpha\in I}$ es relativamente compacta.

Se observa que los procesos estocásticos en la definición anterior no necesariamente están definidos en un mismo espacio de probabilidad.

Teorema 2.2.17. *Sean (E,r) completo y separable y $\{X^\alpha,\alpha\in I\}$ una familia de procesos estocásticos con trayectorias en $\mathcal{D}([0,\infty),E)$. Entonces, $\{X^\alpha,\alpha\in I\}$ es relativamente compacto si, y sólo si, las siguientes dos condiciones se satisfacen:*

(i) *para toda $\eta>0$ y $t\in\mathbb{Q}^+$, existe un compacto $\Gamma_{\eta,t}\subset E$ tal que*

$$\inf_\alpha\mathbf{P}\left(X_t^\alpha\in\Gamma_{\eta,t}^\eta\right)\geq 1-\eta;$$

(ii) *para toda $\eta>0$ y $T>0$, existe $\delta>0$ tal que*

$$\sup_\alpha\mathbf{P}\left(w'(X^\alpha,\delta,T)\geq\eta\right)\leq\eta.$$

Demostración. Primero, se asume que $\{X^\alpha,\alpha\in I\}$ es relativamente compacto. Se recuerda que, para cada $\alpha\in I$, $\mathbf{P}_\alpha(A)=\mathbf{P}\left(X^\alpha\in A\right)$, donde $\mathbf{P}$ es la medida de probabilidad sobre el espacio de probabilidad en el que el proceso X^α está definido. Como $\{X^\alpha,\alpha\in I\}$ es relativamente compacta y (E,r) es completo y separable, $\{\mathbf{P}_\alpha\}_{\alpha\in I}$ es relativamente compacto y $\mathcal{D}([0,\infty),E)$ es completo y separable(ver Teorema 2.2.11). Entonces, por el Teorema 2.1.22, para cada $\eta>0$, existe $K_\eta\subset\mathcal{D}([0,\infty),E)$ tal que

$$\inf_{\alpha\in I}\mathbf{P}\left(X^\alpha\in K_\eta\right)\geq 1-\eta. \tag{2.2.37}$$

Ahora, ya que K_η es compacto, entonces, por el Teorema 2.2.15, para todo racional no negativo, existe un conjunto compacto $\Gamma_{\eta,t} \subset \mathcal{D}([0,\infty), E)$ tal que $x(t) \in \Gamma_{\eta,t}$ para toda $x \in K_\eta$. Esto, junto con (2.2.37), implica

$$\inf_{\alpha\in I} \mathbf{P}\left(X_t^\alpha \in \Gamma_{\eta,t}\right) \geq \inf_{\alpha\in I} \mathbf{P}\left(X^\alpha \in K_\eta\right) \geq 1-\eta.$$

Por otra parte, por el Teorema 2.2.15 (ii), para cada $T>0$, existe $\delta>0$ tal que

$$\sup_{x\in K_\eta} w'(x,\delta,T) < \eta,$$

para toda $\eta>0$. Entonces

$$\sup_{\alpha\in I} \mathbf{P}\left(w'(X^\alpha,\delta,T)\geq\eta\right) \leq \sup_{\alpha\in I} \mathbf{P}\left(X_\alpha \notin K_\eta\right) \leq 1-\inf_{\alpha\in I} \mathbf{P}\left(X_\alpha\in K_\eta\right) \leq \eta.$$

Así, tomando en cuenta que $\Gamma_{\eta,t}^{\eta}$ puede ser reemplazado por $\Gamma_{\eta,t}$ en (i), las condiciones (i) y (ii) se sostienen.

Recíprocamente, sea $\epsilon>0$ y T un entero positivo tal que $\exp\{-T\}<\epsilon/2$. Sea $\delta>0$ tal que (ii) se sostiene con $\eta=\epsilon/4$,

$$\sup_{\alpha\in I} \mathbf{P}\left(w'\left(X^\alpha,\delta,T\right)\geq\frac{\epsilon}{4}\right)\leq\frac{\epsilon}{4}. \tag{2.2.38}$$

Además, sean $m>\delta^{-1}$ y

$$\Gamma=\bigcup_{i=0}^{mT}\Gamma_{\epsilon 2^{-i-2},\frac{i}{m}}.$$

Se define

$$\Gamma^{\epsilon/4} := \left\{x\in E \,\Big|\, \inf_{y\in\Gamma} r\left(x,y\right)<\frac{\epsilon}{4}\right\}.$$

Se observa que la siguiente relación se satisface

$$\left\{X_{\frac{i}{m}}^\alpha \in \Gamma_{\epsilon 2^{-i-2},\frac{i}{m}}\right\} \subset \left\{X_{\frac{i}{m}}^\alpha \in \Gamma^{\epsilon/4}\right\}, \quad i=0,\ldots,mT. \tag{2.2.39}$$

Entonces, por la relación (2.2.39) y la condición (i), se obtiene[13]

$$\begin{aligned}
\inf_{\alpha\in I}\mathbf{P}\left(X_{\frac{i}{m}}^\alpha\in\Gamma^{\epsilon/4}, i=0,1,\ldots,mT\right) &\geq \inf_{\alpha\in I}\left(1-\sum_{i=0}^{mT}\mathbf{P}\left(X_{\frac{i}{m}}^\alpha\notin\Gamma^{\epsilon/4}\right)\right)\\
&= 1-\sup_{\alpha\in I}\sum_{i=0}^{mT}\mathbf{P}\left(X_{\frac{i}{m}}^\alpha\notin\Gamma^{\epsilon/4}\right)\\
&\geq 1-\sup_{\alpha\in I}\sum_{i=0}^{mT}\mathbf{P}\left(X_{\frac{i}{m}}^\alpha\notin\Gamma_{\epsilon 2^{-i-2},\frac{i}{m}}^{\epsilon 2^{-i-2}}\right).
\end{aligned}$$

[13]Se recuerda que $\sup(A-B)=\sup A-\inf B$ y que $\inf(A-B)=\inf A-\sup B$.

Y ya que

$$\begin{aligned}
1-\sup_{\alpha\in I}\sum_{i=0}^{mT}\mathbf{P}\left(X^{\alpha}_{\frac{i}{m}}\notin\Gamma^{\epsilon 2^{-i-2}}_{\epsilon 2^{-i-2},\frac{i}{m}}\right) &= 1-\sup_{\alpha\in I}\sum_{i=0}^{mT}\left(1-\mathbf{P}\left(X^{\alpha}_{\frac{i}{m}}\in\Gamma^{\epsilon 2^{-i-2}}_{\epsilon 2^{-i-2},\frac{i}{m}}\right)\right)\\
&\geq 1-\sum_{i=0}^{mT}\left(1-\inf_{\alpha\in I}\mathbf{P}\left(X^{\alpha}_{\frac{i}{m}}\in\Gamma^{\epsilon 2^{-i-2}}_{\epsilon 2^{-i-2},\frac{i}{m}}\right)\right)\\
&\geq 1-\sum_{i=0}^{mT}\epsilon 2^{-i-2}\\
&\geq 1-\frac{\epsilon}{2},
\end{aligned}$$

se cumple

$$\inf_{\alpha\in I}\mathbf{P}\left(X^{\alpha}_{\frac{i}{m}}\in\Gamma^{\epsilon/4}, i=0,1,\dots,mT\right)\geq 1-\frac{\epsilon}{2}. \tag{2.2.40}$$

Por otro lado, usando la notación dada en el Lema 2.2.13, sea $A=A(\Gamma,\delta)$, la cual, gracias a la compacidad de Γ y al Lema 2.2.13, es relativamente compacto en $\mathcal{D}([0,\infty),E)$. Además, si $x\in\mathcal{D}([0,\infty),E)$ con $w'(x,\delta,T)<\epsilon/4$ y $x(i/m)\in\Gamma^{\epsilon/4}$ para $i=0,\dots,mT$, entonces, por la definición del módulo de continuidad, existe una partición

$$0=t_0<t_1<\cdots<t_{n-1}<T\leq t_n$$

tal que $\min_{1\leq i\leq n}(t_i-t_{i-1})>\delta$ y

$$\max_{1\leq i\leq n}\sup_{s,t\in[t_{i-1},t_i)} r\left(x(s),x(t)\right)<\frac{\epsilon}{2}. \tag{2.2.41}$$

Gracias a la compacidad de Γ, se pueden elegir $y_i\in\Gamma$ tales que

$$r\left(x(i/m),y_i\right)<\frac{\epsilon}{4},\quad i=0,\dots,mT. \tag{2.2.42}$$

Con esto en mente, se define $x'\in A$ por

$$x'(t):=\begin{cases} y_{\lfloor mt_{i-1}\rfloor+1} & \text{si}\quad t_{i-1}\leq t<t_i,\quad i=1,\dots,n-1;\\ \\ y_{\lfloor mt_{n-1}\rfloor+1} & \text{si}\qquad\qquad t\geq t_{n-1}.\end{cases}$$

Entonces, ya que $m^{-1}<\delta$ y $\min_{1\leq i\leq n}(t_i-t_{i-1})>\delta$, para cada $i=1,\dots,n$ se tiene

$$t_{i-1}\leq\frac{\lfloor mt_{i-1}\rfloor+1}{m}\leq t_{i-1}+\frac{1}{m}<t_i.$$

Así, si $t_{i-1}\leq t<t_i\wedge T$ para algún $i\in\{1,\dots,n\}$, (2.2.41) y (2.2.42) permiten deducir

$$r\left(x(t),x'(t)\right)\leq r\left(x(t),x\left(\frac{\lfloor mt_{i-1}\rfloor+1}{m}\right)\right)+r\left(x\left(\frac{\lfloor mt_{i-1}\rfloor+1}{m}\right),y_{\lfloor mt_{i-1}\rfloor+1}\right)<\frac{\epsilon}{2},$$

por lo que

$$\sup_{0\leq t\leq T} r\left(x(t),x'(t)\right)<\frac{\epsilon}{2}.$$

En consecuencia,

$$\begin{aligned}
d(x',x) &\leq \int_0^\infty \exp\{-u\}\left(\sup_{t\geq 0} r\big(x'(t\wedge u), x(t\wedge u)\big)\wedge 1\right) du \\
&= \int_0^T \exp\{-u\}\left(\sup_{0\leq t\leq u} r\big(x'(t), x(t)\big)\wedge 1\right) du \\
&\quad + \int_T^\infty \exp\{-u\}\left(\sup_{0\leq t\leq u} r\big(x'(t), x(t)\big)\wedge 1\right) du \\
&\leq \int_0^\infty \exp\{-u\}\left(\sup_{0\leq t\leq T} r\big(x'(t), x(t)\big)\wedge 1\right) du \\
&\quad + \int_T^\infty \exp\{-u\} du \\
&< \frac{\epsilon}{2} + \exp\{-T\}.
\end{aligned}$$

Y como $\exp\{-T\} < \frac{\epsilon}{2}$,

$$d(x',x) < \frac{\epsilon}{2} + \exp\{-T\} < \epsilon,$$

Por lo tanto, $x \in A^\epsilon = \left\{x \in \mathcal{D}([0,\infty),E) \,\middle|\, \inf_{y\in A} d\,(x,y) < \frac{\epsilon}{4}\right\}$. En particular,

$$\subset \{X^\alpha \in A^\epsilon\} \supset \left\{X^\alpha_{\frac{i}{m}} \in \Gamma^{\epsilon/4}, i = 0, 1, \ldots, mT \text{ y } w'(X^\alpha, \delta, T) < \frac{\epsilon}{4}\right\}$$

para toda $\alpha \in I$, el cual, gracias a (2.2.38) y (2.2.40), justifica

$$\begin{aligned}
\inf_{\alpha\in I} \mathbf{P}\left(X^\alpha \in A^\epsilon\right) &\geq \inf_{\alpha\in I} \mathbf{P}\left(X^\alpha_{\frac{i}{m}} \in \Gamma^{\epsilon/4}, i = 0, 1, \ldots, mT \text{ y } w'(X^\alpha, \delta, T) < \frac{\epsilon}{4}\right) \\
&\geq 1 - \sup_{\alpha\in I}\left(\sum_{i=0}^{mT} \mathbf{P}\left(X^\alpha_{\frac{i}{m}} \notin \Gamma^{\epsilon/4}\right) + \mathbf{P}\left(w'(X^\alpha, \delta, T) \geq \frac{\epsilon}{4}\right)\right) \\
&\geq 1 - \left(\frac{\epsilon}{2} + \frac{\epsilon}{4}\right) \\
&\geq 1 - \epsilon,
\end{aligned}$$

es decir,

$$\inf_{\alpha\in I} \mathbf{P}_\alpha(A^\epsilon) = \inf_{\alpha\in I} \mathbf{P}\left(X^\alpha \in A^\epsilon\right) \geq 1 - \epsilon.$$

Finalmente, como A es compacto, existe $\{x_1, \ldots, x_n\} \subset A$ tal que $A^\epsilon \subset \bigcup_{i=1}^n Bi$, donde B_i es la bola con centro en x_i y radio 2ϵ, entonces, A^ϵ es compacto. El resultado se sigue del Teoema de Prohorov. Por lo tanto, $\{X^\alpha, \alpha \in I\}$ es relativamente compacto.

■

Corolario 2.2.18. *Sea (E, r) completo y separable, y sea $\{X^n, n \geq 1\}$ una sucesión de procesos estocásticos con trayectorias en $\mathcal{D}([0,\infty), E)$. Entonces, la sucesión $\{X^n, n \geq 1\}$ es relativamente compacta si, y sólo si, se cumple*

(i) para cada $\eta > 0$ y racional $t \geq 0$, existe un compacto $\Gamma_{\eta,t} \subset E$ tal que

$$\liminf_{n\to\infty} \mathbf{P}\left(X^n_t \in \Gamma^\eta_{\eta,t}\right) \geq 1 - \eta;$$

(ii) para cada $\eta > 0$ y $T > 0$, existe $\delta > 0$ tal que

$$\limsup_{n\to\infty} \mathbf{P}\left(w'\left(X^n, \delta, T\right) \geq \eta\right) \leq \eta.$$

Demostración. Si la sucesión $\{X^n, n \geq 1\}$ es relativamente compacta, entonces las condiciones (i) y (ii) enunciadas en el corolario se siguen del Teorema 2.2.17. Así, es suficiente probar que si las condiciones (i) y (ii) del corolario se cumplen, entonces la sucesión $\{X^n, n \geq 1\}$ es relativamente compacta. Para ello, sean $T, \eta > 0$ y un racional $t \geq 0$. Por hipótesis, existe un conjunto compacto $\Gamma_0 \in E$, un $\delta_0 > 0$ y un entero positivo n_0 tales que

$$\inf_{n\geq n_0} \mathbf{P}\left(X_t^n \in \Gamma_0^\eta\right) \geq 1-\eta \tag{2.2.43}$$

y

$$\sup_{n\geq n_0} \mathbf{P}\left(w'\left(X^n, \delta_0, T\right) \geq \eta\right) \leq \eta. \tag{2.2.44}$$

Por otra parte, por el Teorema 2.1.22, para cada $n \geq 1$, existe un conjunto compacto $\Gamma_n \subset E$ tal que

$$\mathbf{P}\left(X_t^n \in \Gamma_n^n\right) \geq 1-\eta$$

y, por el Lema 2.2.14 (i), existe $\delta_n > 0$ tal que

$$\mathbf{P}\left(w'\left(X^n, \delta_n, T\right) \geq \eta\right) \leq \eta.$$

Sea $\Gamma = \cup_{n=0}^{n_0-1}\Gamma_i$ y $\delta = \bigwedge_{n=0}^{n_0-1} \delta_n$. Entonces, n_0 puede ser reemplazado por 1 en 2.2.43) y (2.2.44) si se sustituye Γ_0^η por Γ y δ_0 por δ en (2.2.43) y (2.2.44), respectivamente, y conseguir las condiciones (i) y (ii) del Teorema 2.2.17. Por lo tanto, la sucesión $\{X^n, n \geq 1\}$ sea relativamente compacta.

■

Lema 2.2.19. *Si X es un proceso con trayectorias en $\mathcal{D}([0,\infty), E)$, entonces el complemento en $[0,\infty)$ de*

$$\boldsymbol{D}(X) := \left\{t \geq 0 \,\middle|\, \mathbf{P}\left(X_t = X_{t-}\right) = 1\right\} \tag{2.2.45}$$

es a lo más numerable.

Demostración. Para cada ϵ, δ reales positivos y T entero positivo, se define

$$A(\epsilon, \delta, T) := \left\{0 \leq t \leq T \,\middle|\, \mathbf{P}\left(r\left(X_t, X_{t-}\right) \geq \epsilon\right) \geq \delta\right\}.$$

Se afirma que $A(\epsilon, \delta, T)$ es un conjunto finito. Si $A(\epsilon, \delta, T)$ contiene una sucesión infinita $\{t_n\}_{\geq 1}$ de puntos distintos, entonces

$$\begin{aligned}
\mathbf{P}\left(\limsup_{n\to\infty}\left\{r\left(X_{t_n}, X_{t_n-}\right) \geq \epsilon\right\}\right) &= \mathbf{P}\left(\bigcap_{n\geq 1}\bigcup_{m\geq n}\left\{r\left(X_{t_m}, X_{t_m-}\right) \geq \epsilon\right\}\right) \\
&= \lim_{n\to\infty} \mathbf{P}\left(\bigcup_{m\geq n}\left\{r\left(X_{t_m}, X_{t_m-}\right) \geq \epsilon\right\}\right) \\
&\geq \lim_{n\to\infty} \mathbf{P}\left(r\left(X_{t_n}, X_{t_n-}\right) \geq \epsilon\right) \\
&\geq \delta,
\end{aligned}$$

y en consecuencia,

$$\mathbf{P}\left(\limsup_{n\to\infty}\left\{r\left(X_{t_n},X_{t_n-}\right)\geq\epsilon\right\}\right)>0. \tag{2.2.46}$$

Ahora, si se considera la $\sigma-$álgebra cola generada por la sucesión de variables aleatorias $\{X_{t_n}\}_{n\geq 1}$, la cual se denota por τ, entonces

$$\limsup_{n\to\infty}\left\{r\left(X_{t_n},X_{t_n-}\right)\geq\epsilon\right\}\in\tau.$$

Luego, (2.2.46) y la ley 0-1 de Kolmogorov implica

$$\mathbf{P}\left(\limsup_{n\to\infty}\left\{r\left(X_{t_n},X_{t_n-}\right)\geq\epsilon\right\}\right)=1.$$

En otras palabras, la probabilidad del conjunto conformado por los puntos que pertenecen a infinitos conjuntos

$$A_n:=\left\{r\left(X_{t_n},X_{t_n-}\right)\geq\epsilon\right\}$$

es 1. Entonces, para cada $w\in A_n$, existe una sucesión infinita de puntos distintos en $[0,T]$, $\{t_k\}_{k\geq 1}$, tal que $r\left(X_{t_k},X_{t_k-}\right)\geq\epsilon$, para toda $k\geq 1$. Pero esto contradice el hecho de que, para cada $x\in\mathcal{D}([0,\infty),E)$, el número de puntos $t\in[0,T]$ que satisfacen $r\left(x(t),x(t-)\right)\geq\epsilon$ es a lo más finito (ver Corolario 2.2.2). Así, $A(\epsilon,\delta,T)$ es un conjunto finito, y en consecuencia,

$$D(X)^c=\left\{t\geq 0\,\middle|\,\mathbf{P}\left(r\left(X_t,X_{t-}\right)>0\right)>0\right\}=\bigcup_{n\geq 1}\bigcup_{m\geq 1}\bigcup_{N\geq 1}A\left(\frac{1}{n},\frac{1}{m},N\right)$$

es a lo más numerable. ■

Teorema 2.2.20. *Sean E un espacio separable y $X^n, n=1,2,\ldots$, y X procesos estocásticos con trayectorias en $\mathcal{D}([0,\infty),E)$.*

(i) *Si $X^n\Rightarrow X$, entonces*

$$\left(X^n_{t_1},\ldots,X^n_{t_k}\right)\Rightarrow\left(X_{t_1},\ldots,X_{t_k}\right) \tag{2.2.47}$$

para todo conjunto finito $\{t_1,\ldots,t_k\}\subset\boldsymbol{D}(X)$. Más aún, para cada conjunto finito $\{t_1,\ldots,t_k\}\subset[0,\infty)$, existen sucesiones $\{t_1^n\}_{n\geq 1}\subset[t_1,\infty),\ldots,\{t_k^n\}_{n\geq 1}\subset[t_k,\infty)$ que convergen a $t_1,\ldots,t_k$ respectivamente, tales que

$$\left(X^n_{t_1^n},\ldots,X^n_{t_k^n}\right)\Rightarrow\left(X_{t_1},\ldots,X_{t_k}\right).$$

(ii) *Si $\{X^n,n\geq 1\}$ es relativamente compacto y existe un conjunto denso $D\subset[0,\infty)$ tal que (2.2.47) se cumple para todo conjunto finito $\{t_1,\ldots,t_k\}\subset D$, entonces $X^n\Rightarrow X$.*

Demostración. (i) Se asume que $X^n\Rightarrow X$. Entonces, usando el Teorema de representación de Skorokhod (Teorema 2.1.30), existe un espacio de probabilidad, $(\Omega,\mathcal{F},\widehat{\mathbf{P}})$, en el cual estan definidos procesos Y y Y^n, $n = 1, 2, \ldots$, con trayectorias en $\mathcal{D}([0,\infty), E)$ y las mismas distribuciones que X y X^n, $n = 1, 2, \ldots$, respectivamente, y tales que

$$\lim_{n\to\infty} d\left(Y^n, Y\right) = 0 \tag{2.2.48}$$

casi seguramente. Se define la aplicación π_t de $\mathcal{D}([0,\infty), E)$ a E por $\pi_t(x) = x(t)$. Se observa que, para cada $t \in D(X) = D(Y)$, el conjunto

$$A(t) := \left\{x \in \mathcal{D}([0,\infty), E) \,\middle|\, x(t) = x(t-)\right\}$$

satisface

$$\widehat{\mathbf{P}}\left[Y \in A_t\right] = \mathbf{P}\left(X \in A_t\right) = \mathbf{P}\left(X_t = X_{t-}\right) = 1,$$

y en consecuencia, la aplicación π_t es continua casi seguramente respecto a la medida inducida por el proceso Y. Así, extendiendo el razonamiento anterior, si

$$\{t_1, \ldots, t_k\} \subset D(X) = D(Y),$$

entonces, la aplicación $(\pi_{t_1}, \ldots, \pi_{t_k})$ de $\mathcal{D}([0,\infty), E)$ a E^k dada por

$$\left(\pi_{t_1}, \ldots, \pi_{t_k}\right)(x) = \left(x(t_1), \ldots, x(t_k)\right),$$

es continua casi seguramente respecto a la medida inducida por el proceso Y. Entonces, por el Teorema 2.1.15 y (2.2.48), la siguientes igualdades se cumplen casi seguramente

$$\begin{aligned}
\lim_{n\to\infty}\left(Y^n_{t_1}, \ldots, Y^n_{t_k}\right) &= \lim_{n\to\infty}\left(\pi_{t_1}, \ldots, \pi_{t_k}\right)\left(Y^n\right) \\
&= \left(\pi_{t_1}, \ldots, \pi_{t_k}\right)(Y) \\
&= \left(Y_{t_1}, \ldots, Y_{t_k}\right).
\end{aligned}$$

En particular,

$$\lim_{n\to\infty}\left(Y^n_{t_1}, \ldots, Y^n_{t_k}\right) = \left(Y_{t_1}, \ldots, Y_{t_k}\right)$$

en probabilidad. La primera afirmación se sigue aplicando la Proposición 2.1.29.

Para la segunda afirmación, sea $\{t_1, \ldots, t_k\} \subset [0,\infty)$. Gracias al Lema 2.2.19, el complemento de $D(X)$ es a lo más numerable, lo cual permite definir, siempre que $[t_k,\infty) \cap D(X)^c$ sea diferente del vacío,

$$s_k := \inf\{s \in D(X)^c \,|\, t_k < s\}, \quad k = 1, 2 \ldots.$$

Sea $\{t^1_n\}_{n\geq 1}$ la sucesión en $[0,\infty)$ dada por $t^1_n = t_1 + 1/n$, $n = 1, 2, \ldots$. Se observa que si $D(X)^c \subset [0, t_1)$, entonces la sucesión $\{t^1_n\}_{n\geq 1}$ está contenida en $[t_1,\infty) \cap D(X)$ y converge a t_1, cuando $n \to \infty$. En cambio, si $D(X)^c \cap [t_1,\infty)$ es distinto del vacío, entonces, existe una subsucesión $\{t^1_n\}_{n\geq 1}$ contenida en $[t_1,\infty) \cap D(X)$, acotada por s_1 y tal que converge a t_1, cuando $n \to \infty$. Por lo tanto, existe una sucesión $\{t^1_n\}_{n\geq 1}$ en $[t_1,\infty) \cap D(X)$ que converge a t_1, cuando $n \to \infty$. Prosiguiendo con este razonamiento, para cada k, existe una sucesión $\{t^k_n\}_{n\geq 1}$ en $[t_k,\infty) \cap D(X)$ tal que converge a t_k. Entonces, por (2.2.47),

$$\left(X^n_{t^m_1}, \ldots, X^n_{t^m_k}\right) \Rightarrow \left(X_{t^m_1}, \ldots, X_{t^m_k}\right)$$

para toda $m \in \mathbb{N}$ y cuando $n \to \infty$. Como $\{t_n^k\}_{n\geq 1} \in D(X)$, la continuidad por la derecha de X implica que

$$\lim_{m\to\infty} \left(X^n_{t_1^m}, \ldots, X^n_{t_k^m} \right) = (X_{t_1}, \ldots, X_{t_k})$$

casi seguramente, y en particular, en probabilidad. Luego, nuevamente por la Proposición 2.1.29,

$$\left(X^n_{t_1^n}, \ldots, X^n_{t_k^n} \right) \Rightarrow (X_{t_1}, \ldots, X_{t_k}).$$

(ii) Sea $\{\mathbf{P}_n\}_{n\geq 1}$ la sucesión de medidas de probabilidad inducida por $\{X^n, n \geq 1\}$. Como $\{X^n, n \geq 1\}$ es relativamente compacta, toda subsucesión de $\{\mathbf{P}_n\}_{n\geq 1}$ es convergente, y en consecuencia, toda subsucesión de $\{X^n, n \geq 1\}$ tiene una subsucesión que converge en distribución. Así, por el Corolario 2.1.23, es suficiente mostrar que toda subsucesión convergente de $\{X^n, n \geq 1\}$ converge en distribución a X. Para ello, y de ser necesario restringuirse a una subsucesión, se supone que $X_n \Rightarrow Y$. Se demostrará que X y Y tienen la misma distribución.

Sean $\{t_1, \ldots, t_k\} \subset D(Y)$ y $f_1, \ldots, f_k \in \overline{C}(E)$, donde se recuerda que $\overline{C}(E)$ es el espacio de funciones real valuadas continuas y acotadas en E. También, sea $\{t_n^i\}_{n\geq 1}$ en $D \cap [t_i, \infty)$, $i = 1, \ldots, k$, tales que

$$\lim_{n\to\infty} t_n^i = t_i, \quad i = 1, \ldots, k.$$

Se define la función real valuada en E^k, h_k, dada por

$$h_k\left((x_1, \ldots, x_k)\right) = \prod_{i=1}^{k} f_i(x_i).$$

Se observa que, para cada $k \in \mathbb{N}$, h_k pertenece a $\overline{C}(E)$. Esta definición, junto con (2.2.47), implica

$$\begin{aligned}
\lim_{n\to\infty} \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X^n_{t_m^i}\right)\right] &= \lim_{n\to\infty} \mathbf{E}\left[h_k\left(\left(X^n_{t_m^1}, \ldots, X^n_{t_m^k}\right)\right)\right] \\
&= \lim_{n\to\infty} \mathbf{E}\left[h_k\left(\left(X_{t_m^1}, \ldots, X_{t_m^k}\right)\right)\right] \\
&= \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X_{t_m^i}\right)\right],
\end{aligned}$$

para cada $m \geq 1$. Por lo tanto, existen enteros $n_1 < n_2 < \cdots$ tales que

$$\left|\mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X_{t_m^i}\right)\right] - \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X^{n_m}_{t_m^i}\right)\right]\right| < \frac{1}{m}. \tag{2.2.49}$$

Más aún, usando el hecho de que $X_{n_m} \Rightarrow Y$ y $\{t_1, \ldots, t_k\} \subset D(Y)$, (i) justifica

$$\left|\mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X^{n_m}_{t_i}\right)\right] - \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(Y_{t_i}\right)\right]\right| < \frac{1}{m} \tag{2.2.50}$$

Además, por la continuidad por la derecha de X y de X^{n_m}, se tiene

$$\lim_{m\to\infty}\left|\mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X_{t_i}\right)\right] - \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X_{t_m^i}\right)\right]\right| = 0 \tag{2.2.51}$$

y

$$\lim_{m\to\infty}\left|\mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X^{n_m}_{t_i^m}\right)\right] - \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X^{n_m}_{t_i}\right)\right]\right| = 0, \tag{2.2.52}$$

respectivamente. Por otra parte, se observa que

$$\begin{aligned}\left|\mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X_{t_i}\right)\right] - \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(Y_{t_i}\right)\right]\right| \leq{} & \left|\mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X_{t_i}\right)\right] - \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X_{t_m^i}\right)\right]\right| \\ & + \left|\mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X_{t_m^i}\right)\right] - \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X^{n_m}_{t_m^i}\right)\right]\right| \\ & + \left|\mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X^{n_m}_{t_m^i}\right)\right] - \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X^{n_m}_{t_i}\right)\right]\right| \\ & + \left|\mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X^{n_m}_{t_i}\right)\right] - \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(Y_{t_i}\right)\right]\right|.\end{aligned}$$

para cada $m \geq 1$. Entonces, por (2.2.49), (2.2.50), (2.2.51), (2.2.52) y la última desigualdad, se obtiene que

$$\mathbf{E}\left[\prod_{i=1}^{k} f_i\left(X_{t_i}\right)\right] = \mathbf{E}\left[\prod_{i=1}^{k} f_i\left(Y_{t_i}\right)\right] \tag{2.2.53}$$

para toda $f_1, \ldots, f_k \in \overline{C}(E)$ y $\{t_1, \ldots, t_k\} \subset D(Y)$. Más aún, si $\{t_1, \ldots, t_k\} \subset [0, \infty)$, entonces, usando el Lema 2.2.19 como en (2.2.2), existen sucesiones $\{t_n^1\}_{n\geq 1}$ en $D(Y) \cap [0, \infty), \ldots, \{t_n^k\}_{n\geq 1}$ en $D(Y) \cap [0, \infty)$ tales que convergen a $t_1, \ldots, t_k$, respetivamente, lo que permite extender, gracias a la continuidad por la derecha de X y Y, (2.2.53) para todo $\{t_1, \ldots, t_k\} \subset [0, \infty)$. Ahora, se considera los siguientes conjuntos

$$\mathcal{A} := \left\{\bigcap_{k=1}^{n} A_k : A_k \in \left\{\pi_t^{-1}(A) \,\middle|\, t \in [0, \infty) \text{ y } A \subset E \text{ abierto}\right\}\right\}.$$

y

$$\mathcal{H} := \left\{A \in \mathcal{S}_E \,\middle|\, \mathbf{E}\left[\mathbf{1}_{\{X\in A\}}\right] = \mathbf{E}\left[\mathbf{1}_{\{Y\in A\}}\right]\right\}.$$

Es claro que

$$\mathcal{F}'_E = \sigma\left\langle \pi_t \,\middle|\, 0 \leq t < \infty \right\rangle \subset \mathcal{A}.$$

Resulta que $\mathcal{A}$ es un $\pi-$sistema y $\mathcal{H}$ es un $\lambda-$sistema. Para la segunda afirmación, es suficiente con recordar (2.2.53) y observar las siguientes relaciones

$$\mathbf{E}\left[\mathbf{1}_{\{X\in E\}}\right] = \mathbf{E}\left[\mathbf{1}_{\{Y\in E\}}\right],$$

$$\mathbf{E}\left[\mathbf{1}_{\{X\in B/A\}}\right]=\mathbf{E}\left[\mathbf{1}_{\{X\in B\}}\right]-\mathbf{E}\left[\mathbf{1}_{\{X\in A\}}\right]$$

y

$$\mathbf{E}\left[\mathbf{1}_{\{X\in\cup_{n\geq 1}A_n\}}\right]=\lim_{n\to\infty}\mathbf{E}\left[\mathbf{1}_{\{X\in A_n\}}\right]=\lim_{n\to\infty}\mathbf{E}\left[\mathbf{1}_{\{Y\in A_n\}}\right]=\mathbf{E}\left[\mathbf{1}_{\{Y\in\cup_{n\geq 1}A_n\}}\right],$$

para $A, B \in \mathcal{H}$ con $A \subset B$ y $\{A_n\}_{n\geq 1}$ una sucesión creciente en $\mathcal{H}$. Más aún, $\mathcal{A} \subset \mathcal{H}$. En efecto, sea a en $\mathcal{A}$, el cual se sabe que tiene la forma

$$a=\bigcap_{k=1}^{n_a}A_k, \qquad \text{con } A_k\in\left\{\pi_t^{-1}(A)\,\middle|\, t\in[0,\infty) \text{ y } A\subset E \text{ abierto}\right\}.$$

Entonces, por (2.2.53), se obtiene

$$\mathbf{E}\left[\mathbf{1}_{\{X\in a\}}\right]=\mathbf{E}\left[\prod_{k=1}^{n_a}\mathbf{1}_{\{X\in A_k\}}\right]=\mathbf{E}\left[\prod_{k=1}^{n_a}\mathbf{1}_{\{\mathbf{Y}\in\mathbf{A_k}\}}\right]=\mathbf{E}\left[\mathbf{1}_{\{Y\in a\}}\right].$$

En consecuencia, $a \in \mathcal{H}$. Por lo tanto, $\mathcal{A} \subset \mathcal{H}$. Esto, junto al hecho de que $\mathcal{A}$ es un $\pi-$sistema, implica, por el teorema de clases monónotonas aplicado a $\mathcal{A}$, $\sigma\langle\mathcal{A}\rangle \subset \mathcal{H}$. Y como ${\mathcal{F}'}_E \subseteq \sigma\langle\mathcal{A}\rangle$, el Teorema 2.2.12 justifica que

$$\mathcal{F}_E={\mathcal{F}'}_E\subseteq\sigma\langle\mathcal{A}\rangle\subset\mathcal{H},$$

donde se recuerda que $\mathcal{F}_E$ es la $\sigma-$álgebra de Borel en $\mathcal{D}([0,\infty),E)$. Por lo tanto,

$$\mathbf{E}\left[\mathbf{1}_{\{X\in A\}}\right]=\mathbf{E}\left[\mathbf{1}_{\{Y\in A\}}\right]$$

para toda $A \in \mathcal{F}_E$, lo que permite concluir que X y Y tienen la misma distribución, y con ello, obtener el resultado deseado.

■

2.3. Espacio $\mathcal{D}([0,\infty),\mathcal{M}_b^+(\mathbb{R}^+))$

El objetivo de esta sección es enunciar y probar un criterio de tensión para una familia de medidas de probabilidad inducidas por una sucesión de procesos estocásticos con trayectorias en $\mathcal{M}_b^+(\mathbb{R}^+)$, el espacio de las medidas de Borel-Radon no negativas en $\mathbb{R}^+$ que son finitas (ver sección 1.3). Es pertinente aclarar que las justificaciones de las afirmaciones que se presentan previo al criterio escapan a los objetivos del trabajo, por lo que se omiten. Si el lector desea profundizar en el tema se pueden consultar [13, 9, 10], que es la fuente principal de este apartado.

Se recuerda que el espacio de medidas temperadas fue introducido por Iscoe para $E=\mathbb{R}^d$, con el propósito de estudiar resultados sobre procesos con valores en espacios de medidas de Radon no finitas, como por ejemplo la medida de Lebesgue. Para fines del trabajo, interesa contar con un criterio de tensión para procesos con trayectorias en $\mathcal{D}([0,\infty),\mathcal{M}_b^+(\mathbb{R}^+))$. Para ello, se plantea el caso general dado en [9].

Sea $\mathcal{C}_c^\infty(\mathbb{R}^d)$ el espacio de funciones real valuadas en $\mathbb{R}^d$, infinitamente diferenciables y con soporte compacto. Para $p \geq 0$ se define

$$\phi_p(x) := (1 + |x|)^{-p}, \quad x \in \mathbb{R}^d,$$

donde $|\cdot|$ es la norma usual en $\mathbb{R}^d$. Sea

$$K_p(\mathbb{R}^d) = \mathcal{C}_c^\infty(\mathbb{R}^d) \cup \{\phi_p\}.$$

Se denota por $\mathcal{M}_p(\mathbb{R}^d)$ al espacio de medidas Radon no negativas μ en $\mathbb{R}^d$ tales que

$$\int_{\mathbb{R}^d} \phi_p(x) d\mu(dx) < \infty,$$

dotado por la topología $p-$vaga, es decir, la topología que hace continua a los mapeos

$$\mu \mapsto \int_{\mathbb{R}^d} \phi_p(x) d\mu(dx)$$

para toda $\phi \in K_p(\mathbb{R}^d)^+$. Este espacio resulta ser un espacio métrico, separable, completo y no es localmente compacto(ver [10]). Debido a que muchos resultados estan disponibles para espacios que son localmente compacto, la no compacidad local de $\mathcal{M}_p(\mathbb{R}^d)$ plantea el problema de refórmular los conceptos anteriores. Por ejemplo, si $\mathcal{M}_p(\mathbb{R}^d)$ fuera localmente compacto (y entre otras condiciones), entonces existe un proceso de Markov con valores en $\mathcal{M}_p(\mathbb{R}^d)$ (ver Teorema en [10, p. 89]). Resulta que $\mathcal{M}_p(\mathbb{R}^d)$ puede ser visto como un subespacio localmente compacto ([10]). Sea $\dot{\mathbb{R}}^d = \mathbb{R}^d \cup \{\tau\}$, donde τ es un punto aislado. Se define, para $p \geq 0$,

$$\dot{\phi}_p(x) := \begin{cases} \phi_p(x) & \text{si} \quad x \in \mathbb{R}^d; \\ \\ 1 & \text{si} \quad x = \tau, \end{cases}$$

Se denota por $\mathcal{M}_p(\dot{\mathbb{R}}^d)$ al espacio de medidas de Radon no negativas μ en $\dot{\mathbb{R}}^d$ tales que

$$\int_{\mathbb{R}^d} \phi_p d\mu + \mu(\{\tau\}) < \infty,$$

y la topología $p-$vaga en $\mathcal{M}_p(\dot{\mathbb{R}}^d)$ definido de la misma forma que para el espacio $\mathcal{M}_p(\mathbb{R}^d)$, considerando las funciones ϕ en

$$K_p(\dot{\mathbb{R}}^d) = \mathcal{C}_c^\infty(\mathbb{R}^d) \cup \{\dot{\phi}_p\}.$$

Ahora, sea $\mathcal{C}_p(\mathbb{R}^d)$ el espacio de funciones $\phi : \mathbb{R}^d \to \mathbb{R}$ continuas tales que

$$\|\phi\|_p := \sup_{x \in \mathbb{R}^d} \left| \frac{\phi(x)}{\phi_p(x)} \right| < \infty,$$

y $\mathcal{C}_p(\dot{\mathbb{R}}^d)$ el espacio de funciones $\phi : \dot{\mathbb{R}}^d \to \mathbb{R}$ tales que

$$\lim_{|x| \to \infty} |x|^{2p} |\phi(x)| = c$$

y $\phi(\tau) = c$, para algún $c \in \mathbb{R}^+$. Resulta que $\mathcal{C}_p(\mathbb{R}^d)$ es un espacio de Banach con respecto a la norma $\|\cdot\|_p$, que $K_p(\mathbb{R}^d) \subset \mathcal{C}_p(\mathbb{R}^d)$ y $K_p(\dot{\mathbb{R}}^d) \subset \mathcal{C}_p(\dot{\mathbb{R}}^d)$. Se definen

$$\langle \mu, \phi \rangle := \int_{\mathbb{R}^d} \phi d\mu, \quad \mu \in \mathcal{M}_p(\mathbb{R}^d), \quad \phi \in \mathcal{C}_p(\mathbb{R}^d);$$

y

$$\langle \mu, \phi \rangle := \int_{\dot{\mathbb{R}}^d} \phi d\mu, \quad \mu \in \mathcal{M}_p(\dot{\mathbb{R}}^d), \quad \phi \in \mathcal{C}_p(\dot{\mathbb{R}}^d).$$

Ahora, para cada $k \geq 1$ se considera los productos cartesianos $\left(\mathcal{C}_p(\mathbb{R}^d)\right)^k$ y $(\mathcal{M}_p(\mathbb{R}^d))^k$ con la topología producto. De manera análoga se define $(\mathcal{M}_p(\dot{\mathbb{R}}^d))^k$ y $(\mathcal{C}_p(\dot{\mathbb{R}}^d))^k$. Con esto en mente, las conveción anterior se pueden escribir como sigue

$$\langle \boldsymbol{\mu}, \boldsymbol{\phi} \rangle = \sum_{i=1}^{k} \langle \mu_i, \phi_i \rangle ,$$

donde $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k) \in \left(\mathcal{M}_p(\mathbb{R}^d)\right)^k$ y $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_k) \in \left(\mathcal{C}_p(\mathbb{R}^d)\right)^k$. De manera similiar se define $\langle \boldsymbol{\mu}, \boldsymbol{\phi} \rangle$ si se toman $\mu \in (\mathcal{M}_p(\dot{\mathbb{R}}^d))^k$ y $\boldsymbol{\phi} \in (\mathcal{C}_p(\dot{\mathbb{R}}^d))^k$. Una consecuencia de la notación establecida anteriormente es que cualquier $\mu \in \mathcal{M}_p(\mathbb{R}^d)$ está unicamente determinada por

$$\{\langle \mu, \phi \rangle \mid \phi \in K_p(\mathbb{R}^d)\},$$

lo cual sigue siendo válido si se sustituyen $\mathcal{M}_p(\mathbb{R}^d)$ y $K_p(\mathbb{R}^d)$ por $\mathcal{M}_p(\dot{\mathbb{R}}^d)$ y $K_p(\dot{\mathbb{R}}^d)$ respectivamente [9].

Como es usual, se denota por $\mathcal{D}([0,\infty), (\mathcal{M}_p(\dot{\mathbb{R}}^d))^k)$ al espacio de funciones de $\mathbb{R}^d$ en $(\mathcal{M}_p(\dot{\mathbb{R}}^d))^k$ que son continuas por la derecha con límites por la izquierda. A continuación, un criterio de tensión para una familia de medidas de probabilidad definidas en el espacio $\mathcal{D}([0,\infty), (\mathcal{M}_p(\dot{\mathbb{R}}^d))^k)$. Para cada $\phi \in (K_p(\dot{\mathbb{R}}^d))^k$ se considera el siguiente mapeo

$$\Pi_\phi : \mathcal{D}([0,\infty), (\mathcal{M}_p(\dot{\mathbb{R}}^d))^k) \to \mathcal{D}([0,\infty), \mathbb{R})$$

definido por

$$(\Pi_\phi x)(t) = \langle x(t), \phi \rangle , \quad t \geq 0,\ x \in \mathcal{D}([0,\infty), (\mathcal{M}_p(\dot{\mathbb{R}}^d))^k).$$

Se observa que este mapeo está bien definido en el sentido que, para cada trayectoria x en $\mathcal{D}([0,\infty), (\mathcal{M}_p(\dot{\mathbb{R}}^d))^k)$, su imagen bajo Π_ϕ es una trayectoria en $\mathcal{D}([0,\infty), \mathbb{R})$. Más aún, por la definición de la topología, Π_ϕ es un mapeo es continuo. Sean $\mathcal{P}(\mathcal{D}([0,\infty), (\mathcal{M}_p(\dot{\mathbb{R}}^d))^k))$ y $\mathcal{P}(\mathcal{D}([0,\infty), \mathbb{R}))$ los espacios de medidas de probabilidad sobre los espacios de Skorohod respectivos con la topología de convergencia débil. Para cada $\phi \in (K_p(\dot{\mathbb{R}}^d))^k$ se define

$$\tilde{\Pi}_\phi : \mathcal{P}[\mathcal{D}([0,\infty), (\mathcal{M}_p(\dot{\mathbb{R}}^d))^k)] \to \mathcal{P}[\mathcal{D}([0,\infty), \mathbb{R})]$$

definido por

$$(\tilde{\Pi}_\phi \mathbf{P})(B) = \mathbf{P}\left(\Pi_\phi^{-1}(B)\right)$$

con $\mathbf{P} \in \mathcal{P}[\mathcal{D}([0,\infty), (\mathcal{M}_p(\dot{\mathbb{R}}^d))^k)]$ y donde B es un subconjunto de Borel de $\mathcal{D}([0,\infty), \mathbb{R})$. El siguiente resultado juega un papel importante en el trabajo pues proporciona un criterio para determinar tensión de una familia de medidas finitas. La prueba sigue las ideas presentadas en [9].

Teorema 2.3.1. *Una sucesión* $\{\mathbf{P}_n\}_{n\geq 1} \subset \mathcal{P}\left(\mathcal{D}(\mathbb{R}^+, (\mathcal{M}_p(\dot{\mathbb{R}}^d))^k\right)$ *es tensa si, y sólo si, la sucesión* $\{\tilde{\Pi}_\phi \mathbf{P}_n\}_{n\geq 1} \subset \mathcal{P}\left(\mathcal{D}(\mathbb{R}^+, \mathbb{R})\right)$ *es tensa para cada* $\phi \in (K_p(\dot{\mathbb{R}}^d))^k$.

Demostración. Por la continuidad de los mapeos Π_ϕ, la sucesión $\{\tilde{\Pi}_\phi \mathbf{P}_n\}_{n\geq 1}$ es tensa para toda $\phi \in (K_p^\infty(\dot{\mathbb{R}}^d))^k$. Para la suficiencia, por el Teorema 4.6 en [12], es suficiente probar que para cada $T > 0$ y $\epsilon > 0$ existe un compacto $\Gamma_{T,\epsilon} \subset (\mathcal{M}_p(\dot{R}^d))^k$ tal que

$$\mathbf{P}_n\left(\mathcal{D}([0,T], \Gamma_{T,\epsilon})\right) > 1 - \epsilon$$

para toda $n \in \mathbb{N}$. Por hipótesis, existe un compacto $K \subset \mathcal{D}(\mathbb{R}^+, \mathbb{R})$ tal que

$$(\tilde{\Pi}_{\dot{\boldsymbol{\phi}}_p} \mathbf{P}_n)(K) > 1 - \epsilon, \ \forall n \in \mathbb{N}, \tag{2.3.1}$$

donde $\dot{\boldsymbol{\phi}}_p = (\dot{\phi}_p, \ldots, \dot{\phi}_p)$. Luego, por el Teorema 2.2.15, para cada $T > 0$, existe un compacto $\Gamma_T \in \mathbb{R}$ tal que

$$\left\{x(t) \,\middle|\, x \in K\right\} \subset \Gamma_T, \ 0 \leq t \leq T.$$

Esto implica que $K \subset \mathcal{D}([0,T], [-c_T, c_T])$, donde c_T es alguna constante positiva. Sea

$$\Gamma_{T,\epsilon} = \left\{\mu \in \mathcal{M}_p(\dot{\mathbb{R}}^d) \,\middle|\, \left\langle \mu, \dot{\phi}_p \right\rangle \leq c_T\right\}^k,$$

el cual, gracias a [10](página 90) y el teorema de Tikhonov, es un conjuto compacto Entonces,

$$\left\{\boldsymbol{\mu} \in (\mathcal{M}_p(\dot{\mathbb{R}}^d))^k \,\middle|\, \left\langle \boldsymbol{\mu}, \dot{\boldsymbol{\phi}}_p \right\rangle \leq c_T\right\} \subset \Gamma_{T,\epsilon},$$

y así,

$$\Pi^{-1}_{\dot{\boldsymbol{\phi}}_p}(K) \subset \mathcal{D}([0,T], \Gamma_{T,\epsilon}),$$

y en consecuencia, por (2.3.1),

$$\mathbf{P}_n\left(\mathcal{D}([0,T], \Gamma_{T,\epsilon})\right) \geq \mathbf{P}_n\left(\Pi^{-1}_{\dot{\boldsymbol{\phi}}_p}(K)\right) = (\tilde{\Pi}_{\dot{\boldsymbol{\phi}}_p} \mathbf{P}_n)(K) > 1 - \epsilon, \ \forall n \in \mathbb{N}.$$

■

Capítulo 3

El proceso de edades

El proceso de edades está basado en el modelo de edades planteado por primera vez en [15] y consiste en estudiar la variación en el tiempo del tamaño de una población cuyos miembros están caracterizados por su edad considerando las modificaciones producidas por los fenómenos de nacimiento, muerte y envejecimiento de los miembros de la población. Este capítulo tiene como propósito probar un teorema límite de difusión en el sentido de convergencia débil de procesos estocásticos con valores en espacio de medidas siguiendo las ideas presentadas en [6] y [5]. Para lograrlo, se verificará la propiedad de tensión y convergencia de las distribuciones finito dimensionales de los procesos aproximantes.

3.1. Descripción del proceso de edades

Se considera una población donde cada individuo tiene un único descendiente o muere con tasa constante e igual a λ. Sea $\mathbf{N} = (\mathbf{N}_t)_{t\geq 0}$ el tamaño de la población. El proceso $\mathbf{N}$ resulta ser un proceso de nacimiento y muerte, crítico, lineal, sobre los enteros no negativos y con intensidad de salto en el tiempo t dado por $2\lambda\mathbf{N}_t$. Se observa que esta población, al tiempo $t \geq 0$, consiste de $k_j(t,\omega)$ individuos con edad $x_j(t,\omega)$, $0 \leq j \leq J(t,\omega)$, donde $J(t,\omega)$ es el número de los distintos grupos de individuos con la misma edad al tiempo t.

Definición 3.1.1. *Se recuerda que $\mathcal{M}_b^+(\mathbb{R}^+)$ denota el espacio de medidas positivas y finitas definidas en $\mathbb{R}^+$.* ***El proceso de edades*** *(o la distribución de edades) del proceso $\mathbf{N}$, $\mathbf{X} = \{\mathbf{X}_t, t \geq 0\}$, es el proceso estocástico con valores en $\mathcal{M}_b^+(\mathbb{R}^+)$ dado por*

$$\mathbf{X}_t(\cdot) = \sum_{j=1}^{J} k_j \delta_{x_j}(\cdot). \tag{3.1.1}$$

Es importante señalar que la distribución de edades descrita anteriormente no es una cantidad normalizada y que las variables k_j, x_j y J dependen de los parámetros t y ω; sin embargo, esta dependencia no se hará explícito. También, es importante señalar que $\mathbf{X}_t$ puede ser visto como la distribución de frecuencia de las edades al tiempo t. Por ejemplo, para $x_1 \leq x_2$, $\mathbf{X}_t([x_1,x_2])$ representa la cantidad de individuos con edad en el intervalo $[x_1,x_2]$ al tiempo t.
En un nacimiento se agrega una partícula con edad cero a la población y la partícula madre se mantiene viva. En este caso, la expresión (3.1.1) considera, adicionalmente, una

delta de Dirac con masa 1 en 0. En una muerte, asumiendo que la incidencia de muerte es independiente de la edad, se remueve una partícula de la población con probabilidad uniforme, y a la expresión (3.1.1) es modificada por la sustracción de una delta de Dirac con masa 1 en la edad del individuo removido. A continuación, se enlistan los axiomas establecidos en [15] para el modelo de edades de un proceso de nacimiento y muerte crítico con parámetro λ.

1. Las subpoblaciones generadas por dos individuos que coexisten se desarrollan totalmente independientes uno de otros.
2. Un individuo vivo al tiempo t tiene una probabilidad $\lambda h + o(h)$ de producir un nuevo individuo de edad cero durante el intervalo de tiempo posterior de longitud h.
3. Un individuo vivo al tiempo t tiene una probabilidad $\lambda h + o(h)$ de morir durante el intervalo de tiempo posterior de longitud h.
4. Todos los individuos envejecen de forma lineal en el tiempo.

Antes de continuar, se establece la siguiente notación. Para $h > 0$ y algún $n \in \mathbb{N}$, se define el operador traslación O_h por

$$O_h \left(\sum_{j=1}^{n} k_j \delta_{x_j} \right) := \sum_{j=1}^{n} k_j \delta_{x_j+h}.$$

En términos del proceso de edades, $O_h(\mathbf{X}_t)$ aumenta la edad a h unidades a cada partícula viva en la población al tiempo t. Se observa que

$$\mathbf{N}_t = \sum_{j=1}^{J} k_j,$$

o equivalentemente,

$$\mathbf{N}_t = \int_0^\infty \mathbf{X}_t(da).$$

Por lo que, para toda función Borel medible ϕ, se define el siguiente operador

$$\langle \mathbf{X}_t, \phi \rangle = \int_0^\infty \phi(y) \mathbf{X}_t(dy).$$

Entonces, denotando por $\mathbf{1}$ a la indicadora de $[0, \infty)$, se tiene

$$\mathbf{N}_t = \mathbf{X}_t(\mathbb{R}^+) = \langle \mathbf{X}_t, \mathbf{1} \rangle.$$

Ahora bien, los axiomas 1-4 indican que los individuos tienen tiempo de vida exponencial y permiten establecer los siguientes supuestos. Para ello, sea $\mathcal{F}_t = \sigma\langle \mathbf{X}_u, 0 \leq u \leq t \rangle$. Entonces, para h positivo,

1. $\mathbf{P}\left(\mathbf{X}_{t+h} = \mathbf{X}_t + \delta_0 | \mathcal{F}_t\right) = \lambda \langle \mathbf{X}_t, \mathbf{1} \rangle h + o(h)$.
2. $\mathbf{P}\left(\mathbf{X}_{t+h} = \mathbf{X}_t - \delta_y | \mathcal{F}_t\right) = \lambda k_y h + o(h)$, donde k_y denota el número de individuos al tiempo t con edad y.

3. $\mathbf{P}\left(\mathbf{X}_{t+h} = O_h(\mathbf{X}_t)|\mathcal{F}_t\right) = 1 - 2\lambda\langle\mathbf{X}_t, \mathbf{1}\rangle h + o(h)$.

Una consecuencia inmediata de los supuestos anteriores es la siguiente. Si $h > 0$, entonces

$$\begin{aligned}\mathbf{E}\left[\langle\mathbf{X}_{t+h}, \phi\rangle|\mathcal{F}_t\right] &= \langle\mathbf{X}_t + \delta_0, \phi\rangle\langle\mathbf{X}_t, \mathbf{1}\rangle\lambda h + \lambda h\sum_y k_y\langle\mathbf{X}_t - \delta_y, \phi\rangle \\ &\quad + \langle O_h(\mathbf{X}_t), \phi\rangle\left(1 - 2\lambda\langle\mathbf{X}_t, \mathbf{1}\rangle h\right) + o(h) \\ &= \langle\mathbf{X}_t + \delta_0, \phi\rangle\langle\mathbf{X}_t, \mathbf{1}\rangle\lambda h + \langle\mathbf{X}_t, \phi\rangle\Big[\langle\mathbf{X}_t, \mathbf{1}\rangle - 1\Big]\lambda h \\ &\quad + \langle O_h(\mathbf{X}_t), \phi\rangle\left(1 - 2\lambda\langle\mathbf{X}_t, \mathbf{1}\rangle h\right) + o(h),\end{aligned} \tag{3.1.2}$$

donde la suma es sobre todas las edades de los individuos vivos y al tiempo t.

Teorema 3.1.2. ***(Propiedad de Markov.)*** *El proceso* $\mathbf{X}$ *es un proceso de Markov con trayectorias en* $\mathcal{D}\left(\mathbb{R}^+, \mathcal{M}_b^+(\mathbb{R}^+)\right)$, *donde* $\mathcal{D}\left(\mathbb{R}^+, \mathcal{M}_b^+(\mathbb{R}^+)\right)$ *es el espacio de funciones càdlàg definidas en* $\mathbb{R}^+$ *y con valores en* $\mathcal{M}_b^+(\mathbb{R}^+)$, *el espacio de medidas positivas y acotadas en* $\mathbb{R}^+$. *Más aún, el proceso* $\mathbf{X}$ *es un proceso de ramificación.*

Demostración. La propiedad de ramificación se hereda del axioma 1. La propiedad de Markov es consecuencia de la expresión (3.1.2). Primero, se dará la construcción del proceso $\mathbf{X}$. Después, usando un argumento de martingalas, se concluirá que el proceso tiene trayectorias *càdlàg*.
La evolución del tiempo de la distribución de edades puede ser descrito en términos de variables exponenciales y binomiales independientes. Se recuerda que $\mathbf{N}_t$, el tamaño de la población al tiempo t, es igual $\langle\mathbf{X}_t, \mathbf{1}\rangle$. Con esto en mente, se define $\mathbf{J}_0 := 0$ y, para $m \geq 1$,

$$\mathbf{J}_m := \text{ínf}\{t > \mathbf{J}_{m-1}\,\big|\,\langle\mathbf{X}_t, \mathbf{1}\rangle \neq \left\langle\mathbf{X}_{\mathbf{J}_{m-1}}, \mathbf{1}\right\rangle \text{ y } t > \mathbf{J}_{m-1}\} = \text{ínf}\{t\,\big|\,\mathbf{N}_t \neq \mathbf{N}_{\mathbf{J}_{m-1}}\}.$$

Estos tiempos de salto representan el tiempo en el cual sucede un nacimiento o una muerte. Así, por ejemplo, $\mathbf{J}_1$ representa el primer tiempo en el cual sucede el primer nacimiento o la primera muerte. Se observa que la última igualdad permite deducir

$$\mathbf{P}\left(\mathbf{J}_m > s\,\big|\,\mathbf{J}_{m-1}\right) = \exp\{-2\lambda s N_{\mathbf{J}_{m-1}}\} = \exp\{-2\lambda s\left\langle\mathbf{X}_{\mathbf{J}_{m-1}}, \mathbf{1}\right\rangle\}. \tag{3.1.3}$$

Se estudiará la evolución en el tiempo del proceso $\mathbf{X}$ en términos de estos tiempos de salto. Para ello, se asume que el proceso $\mathbf{X}$ comienza con una medida determinística de la forma $\mathbf{X}_0 = \sum_{j=1}^{n_0} k_j\delta_{x_j}$, esto es, se inicia con k_j partículas de edad $x_j \geq 0$, $j = 1, \ldots, n$. Entonces, para $0 \leq t < \mathbf{J}_1$, la población permanece constante y envejece t unidades, es decir,

$$\mathbf{X}_t := O_t(\mathbf{X}_0) = \sum_{j=1}^{n_0} k_j\delta_{x_j+t}, \quad t \in [0, \mathbf{J}_1).$$

Sean τ_i, $i = 1, \ldots, \langle\mathbf{X}_0, \mathbf{1}\rangle$, variables aleatorias independietes e identicamente distribuidos como una variable aleatoria exponencial de parámetro 2λ. Por (3.1.3), se nota que el evento $\{0 \leq t < J_1\}$ puede expresarse como una unión disjunta finita de eventos de la forma

$$\left\{\tau_i \leq t < \tau_\ell\,\big|\,\ell \neq i \text{ y } \ell = 1, \ldots, \langle X_0, \mathbf{1}\rangle\right\}, \quad i = 1, \ldots, \langle X_0\mathbf{1}\rangle.$$

Ahora, para $\mathbf{J}_1 \leq t < \mathbf{J}_2$, la distribución de edades se describe en dos partes dependiendo de si $\mathbf{J}_1$ corresponde al tiempo de un nacimiento, o si corresponde al tiempo de una muerte. En el primer caso, la población envejece $\mathbf{J}_1 + t$ unidades e incrementa una partícula de edad $t - \mathbf{J}_1$ con probabilidad $1/2$. En el segundo caso, la población disminuye una unidad con probabilidad $1/2\,\langle \mathbf{X}_0, \mathbf{1}\rangle$ y el resto envejece $\mathbf{J}_1 + t$ unidades. En otra palabras,

$$\mathbf{X}_t := \begin{cases} O_t(\mathbf{X}_0) + O_{t-\mathbf{J}_1}(\delta_0) & \text{con probabilidad} \quad 1/2, \\ \\ O_t(\mathbf{X}_0) - \delta_{y_j} & \text{con probabilidad} \quad 1/(2\,\langle \mathbf{X}_0, \mathbf{1}\rangle), \\ & j = 1, \ldots, \langle X_0, \mathbf{1}\rangle \end{cases}$$

Es claro que la construcción anterior puede ser extendido para $\mathbf{J}_n \leq t < \mathbf{J}_{n+1}$, $n \geq 2$. Por lo que $\mathbf{X}_t$ puede definirse como un proceso de Markov(de hecho, esto puede deducirse de la expresión (3.1.2)) sobre un espacio de medidas real valuadas. Más aún, debido al hecho de la no explosividad de $\langle \mathbf{X}_t, \mathbf{1}\rangle = \mathbf{N}_t$, $\mathbf{X}_t$ es una medida acotada en $\mathbb{R}^+$.
Ahora, se verificará que el proceso $\mathbf{X}$ tiene trayectorias *càdlàg* usando un argumento de martingalas. Empleando $O_h(\mathbf{X}_t) = \langle \mathbf{X}_t, \phi(\cdot + h)\rangle$ en (3.1.2) se obtiene

$$\begin{aligned} \mathbf{E}\left[\langle \mathbf{X}_{t+h}, \phi\rangle \,|\mathcal{F}_t\right] = &\langle \mathbf{X}_t, \phi\rangle \langle \mathbf{X}_t, \mathbf{1}\rangle \lambda h + \phi(0) \langle \mathbf{X}_t, \mathbf{1}\rangle \lambda h \\ &+ \langle \mathbf{X}_t, \phi\rangle \langle \mathbf{X}_t, \mathbf{1}\rangle \lambda h - \langle \mathbf{X}_t, \phi\rangle \lambda h \\ &+ \langle \mathbf{X}_t, \phi(\cdot + h)\rangle - 2\lambda \langle \mathbf{X}_t, \phi(\cdot + h)\rangle \langle \mathbf{X}_t, \mathbf{1}\rangle h + o(h) \end{aligned}$$

Restando $\langle \mathbf{X}_t, \phi\rangle$ a ambos lados, dividiendo por h y agrupando, se obtiene

$$\begin{aligned} \frac{\mathbf{E}\left[\langle \mathbf{X}_{t+h}, \phi\rangle \,|\mathcal{F}_t\right] - \langle \mathbf{X}_t, \phi\rangle}{h} &= o(1) + 2\langle \mathbf{X}_t, \phi\rangle \langle \mathbf{X}_t, \mathbf{1}\rangle \lambda + \phi(0) \langle \mathbf{X}_t, \mathbf{1}\rangle \lambda - \langle \mathbf{X}_t, \phi\rangle \lambda \\ &\quad + \frac{\langle \mathbf{X}_t, \phi(\cdot + h)\rangle - \langle \mathbf{X}_t, \phi\rangle}{h} - 2\lambda \langle \mathbf{X}_t, \phi(\cdot + h)\rangle \langle \mathbf{X}_t, \mathbf{1}\rangle \\ &= o(1) + 2\langle \mathbf{X}_t, \phi\rangle \langle \mathbf{X}_t, \mathbf{1}\rangle \lambda + \langle \mathbf{X}_t, \phi(0)\mathbf{1}\lambda - \phi\lambda\rangle \\ &\quad + \left\langle \mathbf{X}_t, \frac{\phi(\cdot + h) - \phi(\cdot)}{h}\right\rangle - 2\lambda \langle \mathbf{X}_t, \phi(\cdot + h)\rangle \langle \mathbf{X}_t, \mathbf{1}\rangle . \end{aligned}$$

Luego, haciendo $h \to 0$, la continuidad de ϕ permite escribir la última suma como sigue

$$\lim_{h\to 0} \frac{\mathbf{E}\left[\langle \mathbf{X}_{t+h}, \phi\rangle \,|\mathcal{F}_t\right] - \langle \mathbf{X}_t, \phi\rangle}{h} = \langle \mathbf{X}_t, \phi(0)\mathbf{1}\lambda - \phi\lambda\rangle + \langle \mathbf{X}_t, \phi'\rangle .$$

donde ϕ' denota la derivada de ϕ. En otras palabras, se ha mostrado que

$$\frac{\partial}{\partial t}\mathbf{E}\left[\langle \mathbf{X}_t, \phi\rangle\right] = \langle \mathbf{X}_t, \phi' - \lambda\phi + \lambda\phi(0)\mathbf{1}\rangle = \langle \mathbf{X}_t, L\phi\rangle , \tag{3.1.4}$$

donde $L\phi = \phi' - \lambda\phi + \lambda\phi(0)\mathbf{1}$. En consecuencia, por la Proposición 1.2.11,

$$\langle X_t, \phi\rangle - \langle X_0, \phi\rangle - \int_0^t \langle X_s, L\phi\rangle \, ds$$

es una martingala, por lo que tiene una modificación con trayectorias *càdlàg*. Más aún, para toda ϕ continua y acotada, $\langle X_t, \phi\rangle$ tiene trayectorias *càdlàg*. Por lo tanto, el proceso $\mathbf{X}$ tiene trayectorias en $\mathcal{D}\left(\mathbb{R}^+, \mathcal{M}_b^+(\mathbb{R}^+)\right)$.

■

Sea $\{\mathbf{Y}^n, n \geq 1\}$ una sucesión de variables no negativas en $\mathbb{R}^+$ no necesariamente independientes y $\mathcal{D}\left(\mathbb{R}^+, \mathcal{M}_b^+(\mathbb{R}^+)\right)$ equipado con la topología de Skorokhod. Para cada $n \geq 1$, sea $\{\mathbf{Y}^{j,n}, 1 \leq j \leq n\}$ una familia de copias independientes e identicamente distribuidas (i.i.d.) de $\mathbf{Y}^n$. También, sean $\{\mathbf{X}^{j,n}, 1 \leq j \leq n, n \geq 1\}$, con $\mathbf{X}_0^{j,n} = \delta_{\mathbf{Y}^{j,n}}$, una familia de copias i.i.d. de $\mathbf{X}$ con trayectorias en $\mathcal{D}\left(\mathbb{R}^+, \mathcal{M}_b^+(\mathbb{R}^+)\right)$. Además, sean μ_λ y μ^n las distribuciones definidas por

$$\langle \mu_\lambda, \phi \rangle = \int_0^\infty \phi(y) \lambda \exp\{-\lambda y\} dy$$

y

$$\langle \mu^n, \phi \rangle = \mathbf{E}[\phi(\mathbf{Y}^n)], \tag{3.1.5}$$

respectivamente. Se nota que, para cada $n \in \mathbb{N}$, las variables $\{\mathbf{Y}^{j,n}, 1 \leq j \geq n\}$ sólo influyen en la condición inicial de los procesos $\{\mathbf{X}^{j,n}, 1 \leq j \leq n\}$ y representan la distribución de las edades de los individuos iniciales. Para continuar, sea $\mathbf{Z} = \{\mathbf{Z}^n\}_{n \geq 1}$ una sucesión de procesos con trayectorias en $\mathcal{D}\left(\mathbb{R}^+, \mathcal{M}_b^+(\mathbb{R}^+)\right)$, dado por

$$\langle \mathbf{Z}_t^n, \phi \rangle = \frac{1}{n} \sum_{j=1}^n \langle \mathbf{X}_{nt}^{j,n}, \phi \rangle. \tag{3.1.6}$$

Se observa que, en particular, se tiene

$$\langle \mathbf{Z}_0^n, \phi \rangle = \frac{1}{n} \sum_{j=1}^n \langle \mathbf{X}_0^{j,n}, \phi \rangle = \frac{1}{n} \sum_{j=1}^n \phi(\mathbf{Y}^{j,n}).$$

Recordando que $\overline{\mathbf{N}}$ es la difusión de Feller de parámetro λ dado en el Ejemplo 1.1.4, denotando por $\Rightarrow$ la convergencia débil en la topología de Skorokhod del espacio $\mathcal{D}\left(\mathbb{R}^+, \mathcal{M}_b^+(\mathbb{R}^+)\right)$ y usando la notación $e(x) = \lambda \exp\{-\lambda x\}$, el resultado principal de este capítulo es el siguiente:

Teorema 3.1.3. *Sea $T \geq 0$ y $\overline{\mathbf{N}}$ la difusión de Feller de parámetro λ dada en el ejemplo 1.1.4. Se supone que la sucesión de variables aleatorias $\{\mathbf{Z}_0^n, n \geq 1\}$ tiene un límite en distribución $\overline{\mathbf{Z}}_0$, cuando $n \to \infty$. Entonces, las distribiciones finito dimensionales de $\mathbf{Z}^n$ converge débilmente hacia un proceso $\overline{\mathbf{Z}}$ dado por*

$$\overline{\mathbf{Z}} = \begin{cases} \overline{\mathbf{Z}}_0, & \text{si} \quad t = 0. \\ \\ \overline{\mathbf{N}}_t \mu_\lambda, & \text{si} \quad t > 0, \end{cases}$$

el cual es tal que $\overline{\mathbf{N}}$ es independiente de $\overline{\mathbf{Z}}_0$. Ahora, se supone $\overline{\mathbf{Z}}_0 = \mu_\lambda$ de modo que $\overline{\mathbf{Z}}_t$, con $0 \leq t \leq T$, sea continuo. Se supone también que μ^n tiene una densidad $g_n(x)$ con respecto a la medida de Lebesgue en $\mathbb{R}^+$ y sea $a_n(x) = g_n(x)/e(x)$. Si

(a) $||g_n - e||_\infty \leq C/n$,

(b) $||a_n(\cdot) - a_n(\cdot + ns)||_\infty \leq Cs$, *entonces*

$$\mathbf{Z}^n \Rightarrow \overline{\mathbf{Z}}. \tag{3.1.7}$$

Observación 3.1.4.

- *Si las v.a.* $\mathbf{Y}^{j,n} \sim e(\lambda)$*, entonces, por la ley de grandes números,* $\mathbf{Z}_0^n$ *converge casi seguramente a* μ_λ.
- *El limite de difusión depende de la difusión de Feller* $\overline{\mathbf{N}}$.
- *En el límite de difusión, el tamaño de los grupos de edades se comporta de la misma manera que la población total pero la edad con peso exponencial.*

La demostración de este teorema, como es usual en situaciones en las que se desea probar convergencia débil de procesos y considerando lo desarrollado en el capítulo anterior, tiene dos partes: (1) mostrar la convergencia de las distribuciones finito dimensionales, la cuales caracterizan el proceso límite; (2) probar la tensión de los procesos aproximantes.

3.2. Convergencia de distribuciones finito-dimensionales

Se denota por $\mathcal{C}_b(\mathbb{R}^+)$ al espacio de funciones $f : \mathbb{R}^+ \to \mathbb{R}$ continuas y acotadas, y por $\mathcal{C}_b^+(\mathbb{R}^+)$ el subconjunto de elementos no negativos de $\mathcal{C}_b(\mathbb{R}^+)$. Similarmente, $\mathcal{C}_b^1(\mathbb{R}^+)$ denota el espacio de funciones $f : \mathbb{R}^+ \to \mathbb{R}$ con primera derivada continua y acotada, y por $\mathcal{C}_b^{1,+}(\mathbb{R}^+)$ el subconjunto de elementos no negativos con soporte compacto de $\mathcal{C}_b^1(\mathbb{R}^+)$. Se recuerda que $\mathcal{M}_b^+(\mathbb{R}^+)$ es el espacio de las medidas no negativas y acotadas en $\mathbb{R}^+$ y que el espacio $\mathcal{C}_b(\mathbb{R})$ está dotado de la métrica del supremo, denotada por $\|\cdot\|_\infty$. Para continuar, son necesarios los siguientes resutados preliminares.

Lema 3.2.1. *Para* $\Phi \in \mathcal{C}_b^+(\mathbb{R}^+)$*, sea*

$$\pi(t,\phi,x) := \mathbf{E}_x\left[\exp\left\{\mathbf{i}\left\langle \mathbf{X}_t, \phi\right\rangle\right\}\right] = \mathbf{E}\left[\exp\left\{\mathbf{i}\left\langle \mathbf{X}_t, \phi\right\rangle\right\}|\mathbf{X}_0 = \delta_x\right]$$

el ***funcional característica*** *del proceso* $\mathbf{X}$*. Entonces*

$$\begin{aligned}\pi(t,\phi,x) &= \exp\Big\{\mathbf{i}\phi(x+t)\Big\}\exp\{-2\lambda t\} \\ &\quad + \exp\{-2\lambda t\}\int_0^t \Big[1 + \pi\left(s,\phi,x+t-s\right)\pi\left(s,\phi,0\right)\Big]\lambda\exp\{2\lambda s\}ds.\end{aligned} \tag{3.2.1}$$

Demostración. Primero, se verificará la identidad (3.2.2) utilizando un argumento de renovación. Sean ξ la ley que gobierna el número de descendientes de cada individuo y τ el primer tiempo de salto del proceso $\mathbf{X}$, el cual se sabe, bajo las condiciones establecidas, se distribuye como una variable exponencial de parámetro 2λ. Entonces

$$\begin{aligned}\pi(t,\phi,x) &= \mathbf{E}_x\left[\mathbf{E}_x\left[\exp\left\{\mathbf{i}\left\langle \mathbf{X}_t,\phi\right\rangle\right\}|\tau\right]\right] \\ &= \mathbf{E}_x\Big[\mathbf{E}_x\Big[\exp\left\{\mathbf{i}\left\langle \mathbf{X}_t,\phi\right\rangle\right\}\mathbf{1}_{\{t<\tau\}} + \exp\left\{\mathbf{i}\left\langle \mathbf{X}_t,\phi\right\rangle\right\}\mathbf{1}_{\{t\geq\tau\}}\Big|\tau\Big]\Big] \\ &= \exp\left\{\mathbf{i}\left\langle \delta_{x+t},\phi\right\rangle\right\}\mathbf{P}\left(\tau>t\right) + \mathbf{E}_x\Big[\mathbf{E}_x\Big[\exp\left\{\mathbf{i}\left\langle \mathbf{X}_t,\phi\right\rangle\right\}\mathbf{1}_{\{t\geq\tau\}}\Big|\tau\Big]\Big],\end{aligned}$$

y esto es igual a[1]

$$\exp\Big\{\mathbf{i}\left\langle \delta_{x+t}, \phi\right\rangle\Big\}\exp\{-2\lambda t\}$$
$$+\int_0^t \Big[\mathbf{P}\left(\xi=0\right)+\mathbf{P}\left(\xi=2\right)\mathbf{E}\left[\exp\left\{\mathbf{i}\left\langle \mathbf{X}_{t-u}, \phi\right\rangle\right\}|\mathbf{X}_0=\delta_{x+u}+\delta_0\right]\Big]2\lambda\exp\{-2\lambda u\}du.$$

Pero la propiedad de ramificación implica que la última suma es igual a

$$\exp\Big\{\mathbf{i}\left\langle \delta_{x+t}, \phi\right\rangle\Big\}\exp\{-2\lambda t\}$$
$$+\int_0^t \Big[1/2+1/2\mathbf{E}_{x+u}\left[\exp\left\{\mathbf{i}\left\langle \mathbf{X}_{t-u}, \phi\right\rangle\right\}\right]\mathbf{E}_0\left[\exp\left\{\mathbf{i}\left\langle \mathbf{X}_{t-u}, \phi\right\rangle\right\}\right]\Big]2\lambda\exp\{-2\lambda u\}du,$$

o equivalentemente, igual a

$$\exp\Big\{\mathbf{i}\left\langle \delta_{x+t}, \phi\right\rangle\Big\}\exp\{-2\lambda t\}+\int_0^t \Big[1+\pi(t-u,\phi,x+u)\pi(t-u,\phi,0)\Big]\lambda\exp\{-2\lambda u\}du.$$

Haciendo el cambio de variable $s=t-u$ en la integral, y tomando en cuenta que

$$\left\langle \delta_{x+t}, \phi\right\rangle=\phi(x+t),$$

la identidad (3.2.2) se sigue.

■

Una manera más explícita para $\pi(t,\phi,x)$ se dió en [16], a saber, $\pi(t,\phi,x)$ es igual a

$$1+\frac{\Big(\exp\Big\{\mathbf{i}\phi(x+t)\Big\}-1\Big)\exp\{-\lambda t\}+\int_0^t\Big(\exp\Big\{\mathbf{i}\phi(y)\Big\}-1\Big)\lambda\exp\{-\lambda t\}dy}{1-\int_0^t\Big(\exp\Big\{\mathbf{i}\phi(y)\Big\}-1\Big)\lambda[1+\lambda(t-y)]\exp\{-\lambda t\}dy}. \qquad (3.2.2)$$

Lema 3.2.2. *Sea $\phi\in\mathcal{C}_b^+(\mathbb{R}^+)$ y $T_t\phi(x)=\mathbf{E}_x\left[\langle X_t,\phi\rangle\right]$ para $t\geq 0$.*

(i) *La familia $\{T_t, t\geq 0\}$ es un semigrupo de operadodres lineales en $\mathcal{C}_b^+(\mathbb{R}^+)$ y está dada por*

$$T_t\phi(x)=\phi(x+t)\exp\{-\lambda t\}+\int_0^t\phi(y)\lambda\exp\{-\lambda y\}dy; \qquad (3.2.3)$$

(ii) *El **funcional log-Laplace***

$$K_t\phi(x)=-\log\mathbf{E}_x\left[\exp\left\{-\left\langle\mathbf{X}_t,\phi\right\rangle\right\}\right]$$

define un semigrupo de contracción, es decir,

$$K_{t+s}\phi=K_t(K_s(\phi)),\quad K_0\phi=\phi,$$

y está dada por

$$K_t\phi(x)=\log\left(1-\frac{T_t(1-\exp\{-\phi\})(x)}{1+\lambda\int_0^t T_r(1-\exp\{-\phi\})(0)dr}\right);$$

[1]Se recuerda que el modelo supone que un individuo sólo puede tener, en promedio, un descendiente.

(iii) *Para $0 \leq s \leq t$, la media y el log-Laplace funcional condicional están dadas por*

$$\mathbf{E}\left[\langle \mathbf{X}_t, \phi\rangle \,|\mathcal{F}_s\right] = \langle \mathbf{X}_s, T_{t-s}\phi\rangle\,,$$

$$-\log \mathbf{E}\left[\exp\left\{-\langle \mathbf{X}_t, \phi\rangle\right\} |\mathcal{F}_s\right] = \langle \mathbf{X}_s, K_{t-s}\phi\rangle\,.$$

Demostración. (i) Primero, se obtendrá la identidad (3.2.3). Por (3.2.2), el funcional característico $\pi(t,\phi,x) = \mathbf{E}\left[\exp\left\{\mathbf{i}\langle \mathbf{X}_t, \phi\rangle\right\} |\mathbf{X}_0 = \delta_x\right]$ tiene la forma

$$\pi(t,\phi,x) = 1 + \frac{\left(\exp\left\{\mathbf{i}\phi(x+t)\right\} - 1\right)\exp\{-\lambda t\} + \int_0^t \left(\exp\left\{\mathbf{i}\phi(y)\right\} - 1\right)\lambda\exp\{-\lambda t\}dy}{1 - \int_0^t \left(\exp\left\{\mathbf{i}\phi(y)\right\} - 1\right)\lambda[1+\lambda(t-y)]\exp\{-\lambda t\}dy}.$$

Luego, como

$$T_t\phi(x) = \mathbf{E}\left[\langle \mathbf{X}_t, \phi\rangle \,|\mathbf{X}_0 = \delta_x\right] = \left.\frac{\partial}{\partial(\mathbf{i}\alpha)}\pi(t,\alpha\phi,x)\right|_{\alpha=0},$$

la identidad (3.2.3) se sigue. La propiedad de semigrupo se puede obtener de la propiedad de Markov; sin embargo, para fines de familiarzar al lector, se verificará esta propiedad de manera directa. Sea $t \geq 0$. Debido a la continuidad de las funciones que aparecen en la definición de T_t, se puede inferir que $T_t\phi \in \mathcal{C}_b^+(\mathbb{R}^+)$ para toda $\phi \in \mathcal{C}_b^+(\mathbb{R}^+)$.
Ahora, se afirma que la familia $\{T_t, t \geq 0\}$ cumple con la propiedad de semigrupo. En efecto,

$$\begin{aligned}
T_s\left(T_t\phi\right)(x) &= T_t\phi(x+t)\exp\{-\lambda t\} + \int_0^t T_t\phi(y)\lambda\exp\{-\lambda y\}dy \\
&= \phi(x+s+t)\exp\{-\lambda(t+s)\} + \exp\{-\lambda s\}\int_0^t \phi(y)\lambda\exp\{-\lambda y\}dy \\
&\quad + \int_0^s \phi(y+t)\exp\{-\lambda t\}\lambda\exp\{-\lambda y\}dy \\
&\quad + \int_0^t \phi(y)\lambda\exp\{-\lambda y\}dy\int_0^s \lambda\exp\{-\lambda y\}dy. \\
&= \phi(x+s+t)\exp\{-\lambda(t+s)\} + \exp\{-\lambda s\}\int_0^t \phi(y)\lambda\exp\{-\lambda y\}dy \\
&\quad + \int_0^s \phi(y+t)\exp\{-\lambda(t+y)\}dy + \left(1 - \exp\{-\lambda s\}\right)\int_0^t \phi(y)\lambda\exp\{-\lambda y\}dy \\
&= \phi(x+s+t)\exp\{-\lambda(t+s)\} + \int_t^{t+s} \phi(y)\exp\{-\lambda y\}dy \\
&\quad + \int_0^t \phi(y)\lambda\exp\{-\lambda y\}dy \\
&= \phi(x+s+t)\exp\{-\lambda(t+s)\} + \int_0^{t+s} \phi(y)\exp\{-\lambda y\}dy \\
&= T_{t+s}\phi(x).
\end{aligned}$$

(ii) La propiedad de semigrupo es una característica general de la transformada de Laplace de un proceso de ramificación a tiempo continuo, propiedad que fue presentada en el primer capítulo. Por otro lado, debido a que

$$\begin{aligned} T_t(1-\exp\{-\phi\})(x) &= -\Big(\exp\Big\{\phi(x+t)\Big\}-1\Big)\exp\{-\lambda t\} \\ &\quad - \int_0^t \Big(\exp\Big\{\phi(y)-1\Big\}\exp\{-\lambda t\}\Big)\, g(y)dy, \end{aligned}$$

y que

$$\lambda\int_0^t T_r\Big(1-\exp\{-\phi\}\Big)(0)dr = -\int_0^t \Big(\exp\Big\{\phi(y)\Big\}-1\Big)\,\lambda[1+\lambda(t-y)]\exp\{-\lambda t\}dy,$$

se cumple

$$K_t\phi(x) = \log\left(1-\frac{T_t(1-\exp\{-\phi\})(x)}{1+\lambda\int_0^t T_r(1-\exp\{-\phi\})(0)dr}\right).$$

La propiedad de contracción es consecuencia de la representación anterior. **(iii)** Por la propiedad de Markov del proceso $\mathbf{X}$, las partículas vivas al tiempo s y edad $x(s)$ evolucionan de acuerdo a copias independientes de $\mathbf{X}$ con $\mathbf{X}_0=\delta_{x(s)}$. Esta observación, junto a la propiedad de ramificación de $\mathbf{X}$, implica

$$\begin{aligned} \mathbf{E}\left[\langle \mathbf{X}_t,\phi\rangle \,\middle|\, \mathcal{F}_s\right] &= \mathbf{E}\left[\langle \mathbf{X}_t,\phi\rangle \,\middle|\, \mathbf{X}_s\right] \\ &= \sum_{x_i\in\,\mathfrak{Sop}(\mathbf{X}_s)} \mathbf{E}_{x_i}\left[\langle \mathbf{X}_{t-s},\phi\rangle\right] \\ &= \sum_{x_i\in\,\mathfrak{Sop}(\mathbf{X}_s)} T_{t-s}\phi(x_i) \\ &= \langle \mathbf{X}_s, T_{t-s}\phi\rangle\,, \end{aligned}$$

donde $\mathfrak{Sop}\,(\mathbf{X}_s)$ denota el soporte de la medida $\mathbf{X}_s$. Usando un argumento similar se obtiene

$$-\log \mathbf{E}\left[\exp\left\{-\langle \mathbf{X}_t,\phi\rangle\right\}|\mathcal{F}_s\right] = \langle \mathbf{X}_s, K_{t-s}\phi\rangle\,.$$

■

La convergencia de las distribuciones finito dimensionales se llevará a cabo identificando el límite distribucional de la sucesión $\langle \mathbf{Z}^n_t,\phi\rangle$ via el funcional de Laplace. Para ello se hará uso de la siguiente notación

$$\overline{\phi} := \langle \mu_\lambda,\phi\rangle\,,\quad \ell_t := \frac{\theta}{1+\lambda t\theta}\quad \text{y}\quad S_t\phi(x) := \frac{T_t\big(1-\exp\{-\phi\}\big)(x)}{1+\lambda\int_0^t T_r\big(1-\exp\{-\phi\}\big)(0)dr}.$$

Por la representación de K_t en el Lema 3.2.2, se tiene

$$K_t\phi = -\log(1-S_t\phi).$$

En el siguiente resultado se da una caracterización de las distribuciones finito dimensionales.

Lema 3.2.3. *Sean $0 \leq t_1 \leq \cdots \leq t_m \leq \tau$ y constantes positivas $\theta_1, \ldots, \theta_m$. Entonces,*

(i)

$$-\log \mathbf{E}\left[\exp\left\{-\sum_{k=1}^{m} \theta_k \overline{\mathbf{N}}_{t_k}\right\} \middle| \overline{\mathbf{N}}_0 = 1\right]$$
$$= \ell_{t_1}(\theta_1 + \ell_{t_2-t_1}(\cdots(\theta_{m-1} + \ell_{t_m-t_{m-1}}(\theta_m))\cdots)).$$

(ii) *Además, para ϕ_k, $k = 1, \ldots, m$, en $\mathcal{C}_b^+(\mathbb{R}^+)$,*

$$-\log \mathbf{E}_x\left[\exp\left\{-\sum_{k=1}^{m} \langle \mathbf{X}_{t_k}, \phi_k\rangle\right\}\right]$$
$$= K_{t_1}(\phi_1 + K_{t_2-t_1}(\cdots(\phi_{m-1} + K_{t_m-t_{m-1}}\phi_m)\cdots))(x).$$

(iii) *Para cualquier $n \geq 1$,*

$$-\log \mathbf{E}_x\left[\exp\left\{-\sum_{k=1}^{m} \langle \mathbf{Z}^n_{t_k}, \phi_k\rangle\right\}\right]$$
$$= nK_{nt_1}\left(\frac{\phi_1}{n} + K_{n(t_2-t_1)}\left(\cdots\left(\frac{\phi_{m-1}}{n} + K_{n(t_m-t_{m-1})}\left(\frac{\phi_m}{n}\right)\right)\cdots\right)\right)(x)$$

Demostración. **(i)** Se procederá por inducción sobre m. El caso $m = 1$ se sigue de la transformada de Laplace del proceso de ramificación a tiempo continuo $\overline{\mathbf{N}}$. Para el caso general, se hará uso de la siguiente consecuencia de la propiedad de Markov

$$-\log \mathbf{E}\left[\exp\left\{-\theta \overline{\mathbf{N}}_t\right\} | \sigma\langle \mathbf{N}_s\rangle\right] = -\log \mathbf{E}_{\overline{\mathbf{N}}_s}\left[\exp\left\{-\theta \overline{\mathbf{N}}_{t-s}\right\}\right] = \overline{\mathbf{N}}_s \ell_{t-s}(\theta). \qquad (3.2.4)$$

Entonces, asumiendo válido el caso $m - 1$, la fórmula (3.2.4) conduce a

$$-\log \mathbf{E}_1\left[\exp\left\{-\sum_{k=1}^{m} \theta_k \overline{\mathbf{N}}_{t_k}\right\}\right]$$
$$= -\log \mathbf{E}_1\left[\exp\left\{-\sum_{k=1}^{m-1} \theta_k \overline{\mathbf{N}}_{t_k}\right\} \mathbf{E}_1\left[\exp\left\{-\theta_m \overline{\mathbf{N}}_{t_m}\right\} \middle| \sigma\left\langle \overline{\mathbf{N}}_{t_{m-1}}\right\rangle\right]\right]$$
$$= -\log \mathbf{E}_1\left[\exp\left\{-\sum_{k=1}^{m-1} \theta_k \overline{\mathbf{N}}_{t_k}\right\} \exp\left\{-\overline{\mathbf{N}}_{t_{m-1}} \ell_{t_m-t_{m-1}}(\theta_m)\right\}\right]$$
$$= -\log \mathbf{E}_1\left[\exp\left\{-\sum_{k=1}^{m-2} \theta_k \overline{\mathbf{N}}_{t_k} - \Big(\theta_{m-1} + \ell_{t_m-t_{m-1}}(\theta_m)\Big) N_{t_{m-1}}\right\}\right]$$

Luego, aplicando la hipótesis de inducción,

$$-\log \mathbf{E}_1\left[\exp\left\{-\sum_{k=1}^{m} \theta_k \overline{\mathbf{N}}_{t_k}\right\}\right] = \ell_{t_1}(\theta_1 + \ell_{t_2-t_1}(\cdots(\theta_{m-1} + \ell_{t_m-t_{m-1}}(\theta_m))\cdots)).$$

(ii) Como en el caso anterior, se procederá por inducción sobre m. Es claro que el caso $m=1$ se sostiene por definición(o convención) de K. Se asume que el resultado es válido para $m-1$. Entonces, por Lema 3.2.2 (iii),

$$\begin{aligned}
&-\log \mathbf{E}_x\left[\exp\left\{-\sum_{k=1}^{m}\langle \mathbf{X}_{t_k},\phi_k\rangle\right\}\right]\\
&\quad=-\log \mathbf{E}_x\left[\exp\left\{-\sum_{k=1}^{m-1}\langle \mathbf{X}_{t_k},\phi_k\rangle\right\}\mathbf{E}_x\Big[\exp\{-\langle \mathbf{X}_{t_m},\phi_m\rangle\}\Big|\mathcal{F}_{t_{m-1}}\Big]\right]\\
&\quad=-\log \mathbf{E}_x\left[\exp\left\{-\sum_{k=1}^{m-1}\langle \mathbf{X}_{t_k},\phi_k\rangle\right\}\exp\left\{-\langle \mathbf{X}_{t_{m-1}},K_{t_m-t_{m-1}}\phi_m\rangle\right\}\right]\\
&\quad=-\log \mathbf{E}_x\left[\exp\left\{-\sum_{k=1}^{m-1}\langle \mathbf{X}_{t_k},\phi_k\rangle-\langle \mathbf{X}_{t_{m-1}},\phi_{m-1}+K_{t_m-t_{m-1}}\phi_m\rangle\right\}\right]\\
&\quad=K_{t_1}(\phi_1+K_{t_2-t_1}(\cdots(\phi_{m-1}+K_{t_m-t_{m-1}}\phi_m)\cdots))(x).
\end{aligned}$$

Y el resultado se sigue usando la hipótesis de inducción.
(iii) Primero, se analiza el caso $m=1$. Por definición de $\mathbf{Z}$ y los supuestos de independencia, se tiene que

$$\begin{aligned}
-\log \mathbf{E}_x\left[\exp\left\{-\langle \mathbf{Z}^n_{t_1},\phi\rangle\right\}\right]&=-\log \mathbf{E}_x\left[\exp\left\{-\frac{1}{n}\sum_{j=1}^{n}\left\langle \mathbf{X}^{j,n}_{nt_1},\phi\right\rangle\right\}\right]\\
&=-\sum_{j=1}^{n}\log \mathbf{E}_x\left[\exp\left\{-\frac{1}{n}\left\langle \mathbf{X}^{j,n}_{nt_1},\phi\right\rangle\right\}\right]\\
&=nK_{nt_1}\left(\frac{\phi}{n}\right)
\end{aligned}$$

Siguiendo esta idea, el caso general se sigue de las siguientes identidades

$$-\log \mathbf{E}_x\left[\exp\left\{-\sum_{k=1}^{m}\langle \mathbf{Z}^n_{t_k},\phi_k\rangle\right\}\right]=-\log \mathbf{E}_x\left[\exp\left\{-\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{m}\left\langle \mathbf{X}^{j,n}_{nt_k},\phi_k\right\rangle\right\}\right].$$

Más aún, por la independencia de los procesos $\mathbf{X}^{j,n}$,

$$-\log \mathbf{E}_x\left[\exp\left\{-\sum_{k=1}^{m}\langle \mathbf{Z}^n_{t_k},\phi_k\rangle\right\}\right]=-\sum_{j=1}^{n}\log \mathbf{E}_x\left[\exp\left\{-\sum_{k=1}^{m}\left\langle \mathbf{X}_{nt_k},\frac{\phi_k}{n}\right\rangle\right\}\right].$$

Luego, utilizando (ii) de este enunciado, se obtiene que

$$\begin{aligned}
&-\log \mathbf{E}_x\left[\exp\left\{-\sum_{k=1}^{m}\langle \mathbf{Z}^n_{t_k},\phi_k\rangle\right\}\right]\\
&\quad=\sum_{j=1}^{n}K_{nt_1}\left(\frac{\phi_1}{n}+K_{n(t_2-t_1)}\left(\cdots\left(\frac{\phi_{m-1}}{n}+K_{n(t_m-t_{m-1})}\left(\frac{\phi_m}{n}\right)\right)\cdots\right)\right)(x),
\end{aligned}$$

y con ello lo que se quería mostrar. ■

Es necesario recordar algunas consecuencias del Teorema de Taylor. Para $x \in \mathbb{R}$ tal que $|x| \leq 1/2$ se cumple[2]

$$|\log(1+x) - x| = \left|O\left(|x|^2\right)\right| \leq x^2, \tag{3.2.5}$$

y

$$|1 - \exp\{x\} + x| = \left|O\left(|x|^2\right)\right| \leq x^2. \tag{3.2.6}$$

Lema 3.2.4. *Sea $\phi \in \mathcal{C}_b^+(\mathbb{R}^+)$. Entonces, para $t > 0$ se cumple*

$$\left\|nK_{nt}(\phi/n) - \ell_t(\overline{\phi})\right\|_\infty \to 0, \quad n \to \infty.$$

Demostración. De (3.2.5) y (3.2.6), si $\{h_n\}_{n\geq 1} \subset \mathcal{C}_b^+(\mathbb{R}^+)$ con $||h_n||_\infty \leq C$, entonces, para $n > 2C$,

$$\|-n\log\left(1 - h_n/n\right) - h_n\|_\infty \leq \|h_n\|_\infty^2/n, \tag{3.2.7}$$

y

$$\|1 - \exp\left\{-h_n/n\right\} - h_n/n\|_\infty \leq \|h_n\|_\infty^2/n^2. \tag{3.2.8}$$

Así, tomando $h_n = nS_{nt}(\phi/n)$ y $C = \|\phi\|_\infty$, (3.2.7) implica

$$\begin{aligned}\|nK_{nt}(\phi/n) - nS_{nt}(\phi/n)\|_\infty &= \|-n\log(1 - S_{nt}(\phi/n)) - nS_{nt}(\phi/n)\|_\infty \leq \frac{\|\phi\|^2}{n}\\ &= C^2/n. \end{aligned}\tag{3.2.9}$$

Y como

$$\left\|nK_{nt}(\phi/n) - \ell_t(\overline{\phi})\right\|_\infty \leq \|nK_{nt}(\phi/n) - nS_{nt}(\phi/n)\|_\infty + \left\|nS_{nt}(\phi/n) - \ell_t(\overline{\phi})\right\|_\infty,$$

basta probar que

$$\left\|nS_{nt}(\phi/n) - \ell_t(\overline{\phi})\right\|_\infty \to 0, \quad n \to \infty.$$

Se observa que

$$\left\|nS_{nt}(\phi/n) - \ell_t(\overline{\phi})\right\|_\infty \leq \left\|nS_{nt}(\phi/n) - \overline{\phi}A_n(\phi)^{-1}\right\|_\infty + \left\|\overline{\phi}A_n(\phi)^{-1} - \ell_t(\overline{\phi})\right\|_\infty$$

donde $A_n(\phi) = 1 + \lambda\int_0^t nT_{nr}(1 - \exp\{-\phi/n\})(0)dr$. Pero, debido a que

$$nS_{nt}(\phi/n) = nT_{nt}(1 - \exp\{-\phi/n\})A_n(\phi)^{-1},$$

y $A_n(\phi) > 1$(pues λ es positivo), la última desigualdad se puede escribir como sigue

$$\left\|nS_{nt}(\phi/n) - \ell_t(\overline{\phi})\right\|_\infty \leq \left\|nT_{nt}(1 - \exp\{-\phi/n\}) - \overline{\phi}\right\|_\infty + \left|\overline{\phi}A_n(\phi)^{-1} - \ell_t(\overline{\phi})\right|. \tag{3.2.10}$$

[2]En efecto, si $|x| \leq 1/2$, entonces

$$\left|\sum_{n\geq 2}\frac{x^n}{n}\right| = \left|\sum_{n\geq 0}\frac{x^{n+2}}{n+2}\right| = \left|x^2\sum_{n\geq 0}\frac{x^n}{n+2}\right| \leq |x^2|\sum_{n\geq 0}|x^n| = |x^2|\frac{|x|}{1-|x|} \leq |x^2|.$$

El otro caso en análogo.

Se probará que el lado derecho de (3.2.10) tiende a 0 cuando $n \to \infty$, analizando primero el primer término de el lado derecho, y después, el termino restante. En adelante, se tomará $x \in \mathbb{R}$. Por definición se tiene que $|nT_{nt}(1-\exp\{\phi/n\})(x) - T_{nt}(\phi)(x)|$ es igual a

$$\left| n\left((1-\exp\{-\phi(x+nt)/n\})\exp\{-\lambda nt\} + \int_0^{nt} (1-\exp\{-\phi(y)/n\})\lambda\exp\{-\lambda y\}dy \right)\right.$$

$$\left. -(\phi(x+nt)\})\exp\{-\lambda nt\} - \int_0^{nt} \phi(y)\lambda\exp\{-\lambda y\}dy \right|.$$

O equivalentemente, igual a

$$\left| n\Big(1-\exp\{-\phi(x+nt)/n\} - \phi(x+nt)/n\Big)\exp\{-\lambda nt\} \right.$$
$$\left. + n\int_0^{nt} \Big(1-\exp\{-\phi(y)/n\} - \phi(y)/n\Big)\lambda\exp\{-\lambda y\}dy \right|.$$

Más aún, por (3.2.8), lo expresión anterior es menor o igual que

$$\left| n\left(\frac{\|\phi\|_\infty^2}{n^2}\right)\exp\{-\lambda nt\} + n\int_0^{nt}\left(\frac{\|\phi\|_\infty^2}{n^2}\right)\lambda\exp\{-\lambda y\}dy \right|.$$

En resumen,

$$\Big| nT_{nt}(1-\exp\{\phi/n\})(x) - T_{nt}(\phi)(x) \Big|$$
$$\leq \left| \left(\frac{\|\phi\|_\infty^2}{n}\right)\left(\exp\{-\lambda nt\} + \int_0^{nt}\lambda\exp\{-\lambda y\}dy\right) \right|.$$

Luego, como

$$\left| \left(\frac{\|\phi\|_\infty^2}{n}\right)\left(\exp\{-\lambda nt\} + \int_0^{nt}\lambda\exp\{-\lambda y\}dy\right) \right|$$
$$\leq \left| \left(\frac{C^2}{n}\right)\Big(\exp\{-\lambda nt\} + (1-\exp\{-\lambda nt\})\Big) \right|,$$

$$|nT_{nt}(1-\exp\{\phi/n\})(x) - T_{nt}(\phi)(x)| \leq \frac{C^2}{n},$$

es decir,

$$\|nT_{nt}(1-\exp\{\phi/n\}) - T_{nt}(\phi)\|_\infty \leq \frac{C^2}{n}, \tag{3.2.11}$$

Por otro lado, se cumple

$$\begin{aligned}\left|T_{nt}\phi(x) - \overline{\phi}\right| &= \left|\phi(x+nt)\exp\{-\lambda nt\} + \int_0^{nt} \phi(y)\lambda\exp\{-\lambda y\}dy - \overline{\phi}\right| \\ &= \left|\phi(x+nt)\exp\{-\lambda nt\} - \int_{nt}^{\infty} \phi(y)\lambda\exp\{-\lambda y\}dy\right| \\ &\leq \left|\phi(x+nt)\exp\{-\lambda nt\}\right| + \int_{nt}^{\infty} \left|\phi(y)\lambda\exp\{-\lambda y\}\right| dy \\ &\leq \|\phi\|_\infty \exp\{-\lambda nt\} + \|\phi\|_\infty \left(\exp\{-\lambda nt\} - 1\right) \\ &\leq 2\|\phi\|_\infty \exp\{-\lambda nt\},\end{aligned}$$

esto es,

$$\left\|T_{nt}\phi - \overline{\phi}\right\|_\infty \leq 2\|\phi\|_\infty \exp\{-\lambda nt\}. \tag{3.2.12}$$

Luego, de (3.2.11) y (3.2.12) se tiene

$$\begin{aligned}\left\|nT_{nt}(1-\exp\{-\phi/n\}) - \overline{\phi}\right\|_\infty &\leq \left\|nT_{nt}(1-\exp\{-\phi/n\}) - T_{nt}\phi\right\| + \left\|T_{nt}\phi - \overline{\phi}\right\|_\infty \\ &\leq C^2/n + 2\|\phi\|_\infty \exp\{-\lambda nt\}.\end{aligned}$$

En otras palabras,

$$\left\|nT_{nt}(1-\exp\{-\phi/n\}) - \overline{\phi}\right\|_\infty \leq C^2/n + 2\|\phi\|_\infty \exp\{-\lambda nt\}. \tag{3.2.13}$$

Por otra parte, se observa que $A_n(\phi)$ y $1+\lambda t\overline{\phi}$ no dependen de x. Entonces

$$\begin{aligned}\left|A_n(\phi) - (1+\lambda t\overline{\phi})\right| &= \left|\lambda\left(\int_0^t nT_{nr}(1-\exp\{-\phi/n\})(0)dr - \int_0^t \overline{\phi}dr\right)\right| \\ &\leq \left|\lambda\int_0^t \left(nT_{nr}(1-\exp\{-\phi/n\})(0) - \overline{\phi}\right) dr\right| \\ &\leq \lambda\int_0^t \left|nT_{nr}(1-\exp\{-\phi/n\})(0) - \overline{\phi}\right| dr \\ &\leq \lambda\int_0^t \left\|nT_{nt}(1-\exp\{-\phi/n\}) - \overline{\phi}\right\|_\infty dr,\end{aligned}$$

el cual, por (3.2.13), tiende a 0 cuando $n \to \infty$. En consecuencia, $A_n(\phi)$ converge a $1+\lambda t\overline{\phi}$, cuando $n \to \infty$. Esta última afirmación, junto con (3.2.10) y (3.2.13), implica

$$\left\|nS_{nt}(\phi/n) - \ell_t(\overline{\phi})\right\|_\infty \leq \left\|nT_{nt}(1-\exp\{-\phi/n\}) - \overline{\phi}\right\|_\infty + \left|\overline{\phi}A_n(\phi)^{-1} - \ell_t(\overline{\phi})\right|.$$

El resultado se sigue del hecho de que el lado derecho converge a 0 cuando $n \to \infty$. ■

Lema 3.2.5. *Para una sucesión $\{f_n, n \geq 0\} \subset \mathcal{C}_b^+(\mathbb{R}^+)$, con $C = \sup_n \|f_n\|_\infty$ y $t > 0$, se tiene*

$$\begin{aligned}\|nK_{nt}(f_n/n) - nK_{nt}(f_0/n)\|_\infty &\leq \|T_{nt}(f_n - f_0)\|_\infty \\ &\quad + \lambda C\int_0^t |T_{ns}(f_n - f_0)(0)|ds \\ &\quad + 2C^2(2+\lambda tC)/n, \quad n > 2C.\end{aligned}$$

Demostración. Primero, se observa, para $n > 2C$, que (3.2.9) implica

$$\begin{aligned}\|nK_{nt}(f_n/n) - nK_{nt}(f_0/n)\|_\infty &\le \|nK_{nt}(f_n/n) - nS_{nt}(f_n/n)\|_\infty + \|nK_{nt}(f_0/n) - nS_{nt}(f_0/n)\|_\infty \\ &\quad + \|nS_{nt}(f_n/n) - nS_{nt}(f_0/n)\|_\infty \\ &\le 2C^2/n + \|nS_{nt}(f_n/n) - nS_{nt}(f_0/n)\|_\infty .\end{aligned}$$

Así, es suficiente analizar el segundo término de la última desigualdad. Por definición,

$$\|nS_{nt}(f_n/n) - nS_{nt}(f_0/n)\|_\infty$$

es igual a

$$\left\|nT_{nt}\Big(1 - \exp\{-f_n/n\}\Big)A_n(f_n)^{-1} - nT_{nt}\Big(1 - \exp\{-f_0/n\}\Big)A_n(f_0)^{-1}\right\|_\infty ,$$

el cual, sumando y restando $T_{nr}\Big(n(1 - \exp\{-f_0/n\})\Big)A_n(f_n)^{-1}$ dentro de la norma, es menor o igual que

$$\left\|T_{nt}\Big(n\,(1 - \exp\{-f_n/n\})\Big)A_n(f_n)^{-1} - T_{nt}\Big(n\,(1 - \exp\{-f_0/n\})\Big)A_n(f_n)^{-1}\right\|_\infty$$

$$+\left\|T_{nt}\Big(n\,(1 - \exp\{-f_0/n\})\Big)A_n(f_n)^{-1} - T_{nt}\Big(n\,(1 - \exp\{-f_0/n\})\Big)A_n(f_0)^{-1}\right\|_\infty .$$

Luego, por la linealidad de T_t y recordando que $|A_n(f_n)| \ge 1$ para toda $n \in \mathbb{N}$, la suma anterior es menor o igual que

$$\left\|T_{nt}\Big(n\,(1 - \exp\{-f_n/n\}) - n\,(1 - \exp\{-f_0/n\})\Big)\right\|_\infty$$

$$+\left\|T_{nt}\Big(n\,(1 - \exp\{-f_0/n\})\Big)\Big(A_n(f_n)^{-1} - A_n(f_0)^{-1}\Big)\right\|_\infty ,$$

con A_n como en el Lema 3.2.4. Lo anterior justifica la siguiente desigualdad

$$\begin{aligned}\|nS_{nt}(f_n/n) - nS_{nt}(f_0/n)\|_\infty &\le \left\|T_{nt}\Big(n\,(1 - \exp\{-f_n/n\}) - n\,(1 - \exp\{-f_0/n\})\Big)\right\|_\infty \\ &\quad + \left\|T_{nt}\Big(n\,(1 - \exp\{-f_0/n\})\Big)\right\|_\infty \left|A_n(f_n)^{-1} - A_n(f_0)^{-1}\right|\end{aligned}$$

Con esto en mente, se estudiarán los términos de la derecha de la última desigualdad para probar el enunciado. Para ello, se nota que

$$\left\|T_{nt}\Big(n\,(1 - \exp\{-f_n/n\}) - n\,(1 - \exp\{-f_0/n\})\Big)\right\|_\infty$$

es menor o igual a

$$\left\|T_{nt}\Big(n\log\Big(1 - (1 - \exp\{-f_n/n\})\Big) + n(1 - \exp\{-f_n/n\})\Big)\right\|_\infty$$

$$+\left\|T_{nt}\Big(n\log\Big(1 - (1 - \exp\{-f_0/n\})\Big) + n(1 - \exp\{-f_0/n\})\Big)\right\|_\infty$$

$$+\left\|T_{nt}\Bigg(n\log\Big(1 - (1 - \exp\{-f_0/n\})\Big) - n\log\Big(1 - (1 - \exp\{-f_n/n\})\Big)\Bigg)\right\|_\infty .$$

el cual, usando $h_n = n(1-\exp\{-f_n/n\})$ en (3.2.7) y $h'_n = n(1-\exp\{-f_0/n\})$, es menor o igual que

$$\left\|T_{nt}\left(\|h_n\|_\infty^2/n\right)\right\|_\infty + \left\|T_{nt}\left(\|h'_n\|_\infty^2/n\right)\right\|_\infty + \left\|T_{nt}\Big(f_n - f_0\Big)\right\|_\infty$$

Más aún, ya que $1-\exp\{x\} \le x$ para toda $x \in \mathbb{R}$, la última suma es menor o igual que

$$\left\|T_{nt}\left(\|f_n\|_\infty^2/n\right)\right\|_\infty + \left\|T_{nt}\left(\|f_0\|_\infty^2/n\right)\right\|_\infty + \left\|T_{nt}\Big(f_n - f_0\Big)\right\|_\infty,$$

y como $\|T_{nt}\phi\|_\infty \le \|\phi\|_\infty$,

$$\left\|T_{nt}\left(\|f_n\|_\infty^2/n\right)\right\|_\infty + \left\|T_{nt}\left(\|f_0\|_\infty^2/n\right)\right\|_\infty + \left\|T_{nt}\Big(f_n - f_0\Big)\right\|_\infty = 2C^2/n + \left\|T_{nt}\Big(f_n - f_0\Big)\right\|_\infty,$$

y esto conduce a

$$\left\|T_{nt}\Big(n\left(1-\exp\{-f_n/n\}\right) - n\left(1-\exp\{-f_0/n\}\right)\Big)\right\|_\infty \le 2C^2/n + \left\|T_{nt}\Big(f_n - f_0\Big)\right\|_\infty. \tag{3.2.14}$$

De manera análoga, se tiene que $|A_n(f_n)^{-1} - A_n(f_0)^{-1}|$ es igual a

$$\begin{aligned}
&\left|\lambda\int_0^t nT_{nr}\Big(1-\exp\{-f_0/n\}\Big)(0)dr - \lambda\int_0^t nT_{nr}\Big(1-\exp\{-f_n/n\}\Big)(0)dr\right| \\
&\le \lambda\Bigg(\int_0^t \left|T_{nr}\Big(n\log\Big(1-(1-\exp\{-f_0/n\})\Big) + n(1-\exp\{-f_0/n\})\Big)(0)\right|dr \\
&\quad + \int_0^t \left|T_{nr}\Big(n\log\Big(1-(1-\exp\{-f_n/n\})\Big) + n(1-\exp\{-f_n/n\})\Big)(0)\right|dr \\
&\quad + \int_0^t \left|T_{nr}\Big(f_n - f_0)(0)\right|dr\Bigg) \\
&\le \lambda\Big(2\int_0^t (C^2/n)dr + \int_0^t \left|T_{nr}\Big(f_n - f_0\Big)(0)\right|dr\Big).
\end{aligned}$$

Así,

$$\left|A_n(f_n)^{-1} - A_n(f_0)^{-1}\right| \le 2\lambda t C^2/n + \lambda\int_0^t \left|T_{nr}\Big(f_n - f_0\Big)(0)\right|dr. \tag{3.2.15}$$

Por último, se observa que

$$\|nT_{nt}(1-\exp\{-f_0/n\})\|_\infty \le \|nT_{nt}(f_0/n)\|_\infty \le C.$$

Usando esta desigualdad, junto a (3.2.14) y (3.2.15), en

$$\begin{aligned}
\|nS_{nt}(f_n/n) - nS_{nt}(f_0/n)\|_\infty &\le \left\|T_{nt}\Big(n\left(1-\exp\{-f_n/n\}\right) - n\left(1-\exp\{-f_0/n\}\right)\Big)\right\|_\infty \\
&\quad + \left\|T_{nt}\Big(n\left(1-\exp\{-f_0/n\}\right)\Big)\right\|_\infty \left|A_n(f_n)^{-1} - A_n(f_0)^{-1}\right|,
\end{aligned}$$

se obtiene

$$\begin{aligned}\|nS_{nt}(f_n/n) - nS_{nt}(f_0/n)\|_\infty &\le 2C^2/n + \left\|T_{nt}\Big(f_n - f_0\Big)\right\|_\infty \\ &\quad + C\Big(2\lambda t C^2/n + \lambda\int_0^t \left|T_{nr}\Big(f_n - f_0\Big)(0)\right| dr.\Big).\end{aligned}$$

O equivalentemente,

$$\begin{aligned}\|nS_{nt}(f_n/n) - nS_{nt}(f_0/n)\|_\infty &\le \left\|T_{nt}\Big(f_n - f_0\Big)\right\|_\infty + \lambda C\int_0^t \left|T_{nr}\Big(f_n - f_0\Big)(0)\right| dr \\ &\quad + 2C^2(1+\lambda t C)/n. \end{aligned} \tag{3.2.16}$$

La prueba se sigue usando la desigualdad (3.2.16) en

$$\|nK_{nt}(f_n/n) - nK_{nt}(f_0/n)\|_\infty \le 2C^2/n + \|nS_{nt}(f_n/n) - nS_{nt}(f_0/n)\|_\infty .$$

que se dió al principio de la demostración. ■

Lema 3.2.6. *Para toda $0 = t_0 < t_1 < \cdots < t_m \le \tau$, y $\phi_k \in \mathcal{C}_b^+(\mathbb{R}^+)$,, $k = 1, \ldots, m$, se tiene*

$$\sup_x \left| \log \mathbf{E}_x \left[\exp\left\{ -\sum_{k=1}^m \langle \mathbf{Z}^n_{t_k}, \phi_k\rangle \right\}\right] - \log \mathbf{E}\left[\exp\left\{-\sum_{k=1}^m \phi_k \overline{N}_t\right\}\middle| \overline{N}_0 = 1\right]\right| \to 0,$$

cuando $n \to \infty$.

Demostración. La prueba se hará por inducción sobre m. Para cada $m \in \mathbb{N}$ se considera la siguiente notación

$$L^m(t_0, t_1, \ldots, t_m) = \ell_{t_1-t_0}\left(\overline{\phi}_1 + \ell_{t_2-t_1}\left(\cdots\left(\overline{\phi}_{m-1} + \ell_{t_m-t_{m-1}}\left(\overline{\phi}_m\right)\right)\cdots\right)\right),$$

y

$$\begin{aligned}&H^m(t_0, t_1, \ldots, t_m) \\ &\quad = nK_{n(t_1-t_0)}\left(\frac{\phi_1}{n} + K_{n(t_2-t_1)}\left(\cdots\left(\frac{\phi_{m-1}}{n} + K_{n(t_m-t_{m-1})}\left(\frac{\phi_m}{n}\right)\right)\cdots\right)\right)(x).\end{aligned}$$

El superíndice m denota el número de composiciones que se tienen en el lado derecho de las variables L^m y H^m, que son $m+1$; mientras que $(t_0, t_1, \ldots, t_m)$ determina qué variable t aparece. Por ejemplo,

$$L^{m-1}(t_1, \ldots, t_m) = \ell_{t_2-t_1}\left(\overline{\phi}_2 + \ell_{t_3-t_4}\left(\cdots\left(\overline{\phi}_{m-1} + \ell_{t_m-t_{m-1}}\left(\overline{\phi}_m\right)\right)\cdots\right)\right),$$

De acuerdo al Lema 3.2.3, lo que se quiere probar es

$$\|H^m(t_0, t_1, \ldots, t_m) - L^m(t_0, t_1, \ldots, t_m)\|_\infty \to 0, \quad n \to \infty.$$

El caso $m = 1$ se desarrolló en el Lema 3.2.4. Ahora, se supone válido el enunciado para $m-1$, esto es, se cumple

$$\left\|H^{m-1}(t_1, \ldots, t_m) - L^{m-1}(t_1, \ldots, t_m)\right\|_\infty \to 0, \quad n \to \infty.$$

Ahora, sea $\phi = \phi_1 + L^{m-1}(t_1, t_2, \ldots, t_m)$. Entonces, las siguientes identidades se satisfacen

$$\begin{aligned} \ell_{t_1}(\overline{\phi}) &= \ell_{t_1}(\overline{\phi_1} + L^{m-1}(t_1, t_2, \ldots, t_m)) = L^m(t_0, t_1, \ldots, t_m), \\ nK_{nt_1}\left(\Big(\phi_1 + H_n^{m-1}(t_1, \ldots, t_m)\Big)/n\right) &= H_n^m(t_0, t_1, \ldots, t_m) \end{aligned}$$

Estas consideraciones permiten deducir la siguiente desigualdad

$$\begin{aligned} &\|H_n^m(t_0, t_1, \ldots, t_m) - L^m(t_0, t_1, \ldots, t_m)\|_\infty \\ &\qquad \leq \left\|nK_{nt_1}\left(\Big(\phi_1 + H_n^{m-1}(t_1, \ldots, t_m)\Big)/n\right) - nK_{nt_1}(\phi/n)\right\|_\infty \\ &\qquad + \left\|nK_{nt_1}(\phi/n) - \ell_{t_1}(\overline{\phi})\right\|_\infty, \end{aligned}$$

donde el último término de la derecha converge a cero cuando $n \to \infty$ por el Lema 3.2.4. Basta verificar que el primer término de la derecha en la desigualdad de arriba también converge a cero cuando $n \to \infty$. Sea $C = \sum_{j=1}^m \|\phi_j\|_\infty$. Entonces, para $n > 2C$, se cumple

$$\|\log\{1 \pm \phi_j/n\}\|_\infty \leq \|\phi_j/n\|_\infty .$$

para toda $j = 1, \ldots, m$. Más aún, por esta misma razón se cumple

$$\left\|K_{n(t_j - t_{j-1})}(\phi_j/n)\right\|_\infty \leq \|\phi_j/n\|_\infty ,$$

para $j = 1, \ldots, m$. Usando este hecho se obtiene

$$\left\|\phi_{j-1}/n + K_{n(t_j - t_{j-1})}(\phi_j/n)\right\|_\infty \leq \|\phi_{j-1}/n\|_\infty + \|\phi_j/n\|_\infty ,$$

para $j = 2, \ldots, m$. Continuando con este razonamiento, se consigue

$$\|H^m(t_0, t_1, \ldots, t_m)\|_\infty \leq \sum_{j=1}^m \|\phi_j\|_\infty = C.$$

Así, tomando $f_n = \phi_1 + H_n^m(t_0, t_1, \ldots, t_m)$, $f_0 = \phi$ y C como antes, el Lema 3.2.5 implica, para $n > 2c$,

$$\begin{aligned} &\left\|nK_{nt_1}\left(\Big(\phi_1 + H_n^{m-1}(t_1, \ldots, t_m)\Big)/n\right) - nK_{nt_1}(\phi/n)\right\|_\infty \\ &\qquad \leq 2\left\|T_{nt}\Big(H_n^{m-1}(t_1, \ldots, t_m) - L^{m-1}(t_1, \ldots, t_m)\Big)\right\|_\infty \\ &\qquad + \lambda C \int_0^t \left|T_{ns}\left(H_n^{m-1}(t_1, \ldots, t_m) - L^{m-1}(t_1, \ldots, t_m)\right)(0)\right| ds \\ &\qquad + 2C^2(2 + \lambda t C)/n. \end{aligned}$$

Luego, por la propiedad de contracción de T_t, la suma anterior es menor o igual que

$$(1 + \lambda C t)\left\|H_n^{m-1}(t_1, \ldots, t_m) - L^{m-1}(t_1, \ldots, t_m)\right\|_\infty + 2C^2(2 + \lambda t C)/n,$$

el cual, por hipótesis de inducción, converge a cero cuando $n \to \infty$. Por lo tanto,

$$\left\|nK_{nt_1}\left(\Big(\phi_1 + H_n^{m-1}(t_1, \ldots, t_m)\Big)/n\right) - nK_{nt_1}(\phi/n)\right\|_\infty$$

converge a cero cuando $n \to \infty$, y en consecuencia,

$$\|H^m(t_0, t_1, \ldots, t_m) - L^m(t_0, t_1, \ldots, t_m)\|_\infty \to 0, \quad n \to \infty.$$

$\blacksquare$

Con lo desarrollado hasta ahora, se puede demostrar la convergencia de las distribuciones finito dimensionales del Teorema 3.1.3. La siguinte proposición verifica esta covergencia.

Proposición 3.2.7. *Para toda* $0 = t_0 < t_1 < \cdots < t_m \leq \tau$ *se tiene*

$$(\mathbf{Z}_0^n, \mathbf{Z}_{t_1}^n, \ldots, \mathbf{Z}_{t_m}^n) \Rightarrow (\overline{\mathbf{Z}_0}, \overline{\mathbf{N}}_{t_1}\mu_\lambda, \ldots, \overline{\mathbf{N}}_{t_k}\mu_\lambda), \quad n \to \infty,$$

donde el proceso de ramificación continuo $\overline{\mathbf{N}}$*, con* $\overline{\mathbf{N}}_0 = 1$*, es independiente de* $\overline{\mathbf{Z}}_0$*.*

Demostración. Se procederá por inducción sobre m. Sean $\phi_k \in \mathcal{C}_b^+(\mathbb{R}^+)$, $k = 0, \ldots, m$. Se quiere probar que

$$\mathbf{E}\left[\exp\left\{-\sum_{k=0}^{m}\left\langle \mathbf{Z}_{t_k}^n, \phi_k\right\rangle\right\}\right]$$

converge, cuando $n \to \infty$, hacia

$$\mathbf{E}\left[\exp\{-\left\langle\overline{\mathbf{Z}}_0, \phi_0\right\rangle\}\right]\mathbf{E}\left[\exp\left\{-\sum_{k=0}^{m}\overline{\phi}_k\overline{\mathbf{N}}_{t_k}\right\}\right].$$

Se observa que el caso $m = 0$ corresponde a la hipótesis de que $\mathbf{Z}_n^0$ tiene un límite en distribución $\overline{\mathbf{Z}}_0$, cuando $n \to \infty$. Primero, se analiza el caso $m = 1$. Para ello, se considera la sigma álgebra generada por $\mathbf{X}_0$, el cual se denota por σ_0. Entonces, para $t > 0$, la independencia de los procesos implica

$$\begin{aligned}
\mathbf{E}\left[\exp\{-\langle\mathbf{Z}_0^n, \phi_0\rangle - \langle\mathbf{Z}_t^n, \phi_1\rangle\}\right] &= \mathbf{E}\left[\exp\{-\langle\mathbf{Z}_0^n, \phi_0\rangle\}\,\mathbf{E}\left[\exp\{-\langle\mathbf{Z}_t^n, \phi_1\rangle\}\,\Big|\mathcal{F}_0\right]\right]\\
&= \mathbf{E}\left[\exp\{-\langle\mathbf{Z}_0^n, \phi_0\rangle\}\,\mathbf{E}\left[\exp\left\{-\frac{1}{n}\sum_{j=1}^{n}\langle\mathbf{X}_{nt}^{j,n}, \phi_1\rangle\right\}\Big|\mathcal{F}_0\right]\right]\\
&= \mathbf{E}\left[\exp\{-\langle\mathbf{Z}_0^n, \phi_0\rangle\}\prod_{j=1}^{n}\mathbf{E}\left[\exp\left\{-\frac{1}{n}\langle\mathbf{X}_{nt}^{j,n}, \phi_1\rangle\right\}\Big|\mathcal{F}_0\right]\right],
\end{aligned}$$

y este a su vez, es igual a

$$\mathbf{E}\left[\exp\{-\langle\mathbf{Z}_0^n, \phi_0\rangle\}\exp\left\{\sum_{j=1}^{n}\left(-\log\mathbf{E}\left[\exp\left\{-\frac{1}{n}\langle\mathbf{X}_{nt}^{j,n}, \phi_1\rangle\right\}\Big|\mathcal{F}_0\right]\right)\right\}\right].$$

Luego, por los Lemas 3.2.2 y 3.2.3 (iii), la expresión anterior coincide con

$$\mathbf{E}\left[\exp\{-\langle\mathbf{Z}_0^n, \phi_0\rangle\}\exp\left\{-\sum_{j=1}^{n}\langle\mathbf{X}_0^{j,n}, K_{nt}\phi_1\rangle\right\}\right].$$

Y como

$$\begin{aligned}\langle \mathbf{X}_0^{j,n}, K_{nt}\phi_1\rangle &= \int_0^\infty K_{nt}\phi_1(y)\mathbf{X}_0^{j,n}(dy)\\ &= n\int_0^\infty K_{nt}\left(\frac{\phi_1}{n}\right)(y)\delta_{\mathbf{Y}_0^{j,n}}(dy)\\ &= nK_{nt}\left(\frac{\phi_1}{n}\right)(\mathbf{Y}_0^{j,n}),\end{aligned}$$

la última esperanza es igual a

$$\mathbf{E}\left[\exp\left\{-\left\langle \mathbf{Z}_0^n,\phi_0\right\rangle\right\}\exp\left\{-\sum_{j=1}^n nK_{nt}\left(\frac{\phi_1}{n}\right)(\mathbf{Y}_0^{j,n})\right\}\right].$$

Por lo que

$$\begin{aligned}\mathbf{E}\left[\exp\{-\left\langle \mathbf{Z}_0^n,\phi_0\right\rangle-\left\langle \mathbf{Z}_t^n,\phi_1\right\rangle\}\right] &= \mathbf{E}\left[\exp\left\{-\left\langle \mathbf{Z}_0^n,\phi_0\right\rangle\right\}\exp\left\{-\sum_{j=1}^n nK_{nt}\left(\frac{\phi_1}{n}\right)(\mathbf{Y}_0^{j,n})\right\}\right]\\ &= \mathbf{E}\left[\exp\left\{-\left\langle \mathbf{Z}_0^n,\phi_0\right\rangle\right\}\exp\left\{\left\langle \mathbf{Z}_0^n, nK_{nt}\left(\frac{\phi_1}{n}\right)\right\rangle\right\}\right].\end{aligned}$$

O equivalentemente,

$$\mathbf{E}\left[\exp\{-\left\langle \mathbf{Z}_0^n,\phi_0\right\rangle-\left\langle \mathbf{Z}_t^n,\phi_1\right\rangle\}\right] = \mathbf{E}\left[\exp\left\{-\left\langle \mathbf{Z}_0^n,\phi_0+nK_{nt}\left(\frac{\phi_1}{n}\right)\right\rangle\right\}\right] \tag{3.2.18}$$

Por otra parte, debido a que $\ell_t(\overline{\phi_1})$ no depende de la edad inicial x y $\left\langle\overline{\mathbf{Z}}_0,\mathbf{1}\right\rangle=1$,

$$\begin{aligned}\mathbf{E}\left[\exp\left\{-\left\langle\overline{\mathbf{Z}}_0,\phi_0\right\rangle\right\}\right]\mathbf{E}\left[\exp\left\{-\overline{\mathbf{N}}_t\overline{\phi}_1\right\}\right] &= \mathbf{E}\left[\exp\left\{-\left\langle\overline{\mathbf{Z}}_0,\phi_0\right\rangle\right\}\right]\exp\left\{-\ell_t(\overline{\phi}_1)\right\}\\ &= \mathbf{E}\left[\exp\left\{-\left\langle\overline{\mathbf{Z}}_0,\phi_0\right\rangle-\left\langle\overline{\mathbf{Z}}_0,\mathbf{1}\right\rangle\ell_t(\overline{\phi}_1)\right\}\right]\\ &= \mathbf{E}\left[\exp\left\{-\left\langle\overline{\mathbf{Z}}_0,\phi_0+\ell_t(\overline{\phi}_1)\right\rangle\right\}\right].\end{aligned}$$

Así, usando (3.2.18) y la última identidad, se obtiene que

$$\left|\mathbf{E}\left[\exp\{-\left\langle \mathbf{Z}_0^n,\phi_0\right\rangle-\left\langle \mathbf{Z}_t^n,\phi_1\right\rangle\}\right]-\mathbf{E}\left[\exp\left\{-\left\langle\overline{\mathbf{Z}}_0,\phi_0\right\rangle\right\}\right]\mathbf{E}\left[\exp\left\{-\overline{\mathbf{N}}_t\overline{\phi}_1\right\}\right]\right|$$

es igual a

$$\left|\mathbf{E}\left[\exp\left\{-\left\langle \mathbf{Z}_0^n,\phi_0+nK_{nt}\left(\frac{\phi_1}{n}\right)\right\rangle\right\}\right]-\mathbf{E}\left[\exp\left\{-\left\langle\overline{\mathbf{Z}}_0,\phi_0+\ell_t(\overline{\phi}_1)\right\rangle\right\}\right]\right|.$$

Luego, se tienen las siguientes desigualdades

$$\begin{aligned}&\left|\mathbf{E}\left[\exp\left\{-\left\langle \mathbf{Z}_0^n,\phi_0+nK_{nt}\left(\frac{\phi_1}{n}\right)\right\rangle\right\}\right]-\mathbf{E}\left[\exp\left\{-\left\langle\overline{\mathbf{Z}}_0,\phi_0+\ell_t(\overline{\phi}_1)\right\rangle\right\}\right]\right|\\ &\le \left|\mathbf{E}\left[\exp\left\{-\left\langle \mathbf{Z}_0^n,\phi_0+nK_{nt}\left(\frac{\phi_1}{n}\right)\right\rangle\right\}\right]-\mathbf{E}\left[\exp\left\{-\left\langle\overline{\mathbf{Z}}_0,\phi_0+\ell_t(\overline{\phi}_1)\right\rangle\right\}\right]\right|\\ &\quad+\left|\mathbf{E}\left[\exp\left\{-\left\langle\mathbf{Z}_0,\phi_0+\ell_t(\overline{\phi}_1)\right\rangle\right\}\right]-\mathbf{E}\left[\exp\left\{-\left\langle\mathbf{Z}_0,\phi_0+\ell_t(\overline{\phi}_1)\right\rangle\right\}\right]\right|\\ &\le \left|\mathbf{E}\left[\exp\left\{-\left\langle \mathbf{Z}_0^n,\phi_0+nK_{nt}\left(\frac{\phi_1}{n}\right)\right\rangle\right\}\right]-\mathbf{E}\left[\exp\left\{-\left\langle\mathbf{Z}_0,\phi_0+\ell_t(\overline{\phi}_1)\right\rangle\right\}\right]\right|\\ &\quad+\left|\mathbf{E}\left[\exp\left\{-\left\langle\mathbf{Z}_0,\phi_0\right\rangle\right\}\right]-\mathbf{E}\left[\exp\left\{-\left\langle\overline{\mathbf{Z}}_0,\phi_0\right\rangle\right\}\right]\right|\exp\left\{-\ell_t(\overline{\phi}_1)\right\}.\end{aligned}$$

Claramente, el último término converge a cero cuando $n \to \infty$ por hipótesis; mientras que

$$\left| \mathbf{E}\left[\exp\left\{ -\left\langle \mathbf{Z}_0^n, \phi_0 + nK_{nt}\left(\frac{\phi_1}{n}\right) \right\rangle \right\} \right] - \mathbf{E}\left[\exp\left\{ -\left\langle \mathbf{Z}_0, \phi_0 + \ell_t(\overline{\phi}_1) \right\rangle \right\} \right] \right|$$

es igual al producto entre

$$\left| \mathbf{E}\left[\exp\left\{ -\left\langle \mathbf{Z}_0^n, \phi_0 + nK_{nt}\left(\frac{\phi_1}{n}\right) \right\rangle \right\} \right] \right|$$

y

$$\left| 1 - \mathbf{E}\left[\exp\left\{ \left\langle \mathbf{Z}_0^n, \phi_0 + nK_{nt}\left(\frac{\phi_1}{n}\right) \right\rangle - \left\langle \mathbf{Z}_0, \phi_0 + \ell_t(\overline{\phi}_1) \right\rangle \right\} \right] \right|.$$

Y, dado que $\exp\{x\} \leq 1$ si $x \leq 0$ y $1 - \exp\{x\} \leq -x$ para toda $x \in \mathbb{R}$, este producto es menor o igual a

$$\mathbf{E}\left[\left| \left\langle \mathbf{Z}_0^n, nK_{nt}\left(\frac{\phi_1}{n}\right) - \ell_t(\overline{\phi}_1) \right\rangle \right| \right] \leq \left\| nK_{nt}(\phi_1/n) - \ell_t(\overline{\phi}_1) \right\|_\infty ,$$

el cual converge a cero cuando $n \to \infty$ por el Lema 3.2.10. En consecuencia, el resultado es válido para el caso $m = 1$. Para el caso $m \geq 2$, las las funciones $nK_{nt}(\phi/n)$ y $\ell_t(\overline{\phi}_1)$ son reemplazadas por las funciones iteradas correspondientes que aparecen en el Lema 3.2.3. Respecto a la estimación, la norma

$$\left\| nK_{nt}(\phi_1/n) - \ell_t(\overline{\phi}_1) \right\|_\infty$$

es sustituida por

$$\| H_n^m(t_0, t_1, \ldots, t_m) - L^m(t_0, t_1, \ldots, t_m) \|_\infty ,$$

el cual se sabe que converge a cero cuando $n \to \infty$ por el Lema 3.2.6. ∎

3.3. Tensión

Para establecer la propiedad de tensión de la sucesión $\{\mathbf{Z}^n\}$ requerido en el Teorema 3.1.3, se hará uso de un resultado establecido en el Teorema 2.3.1. En términos del proceso de edades, lo anterior se traduce como sigue.

Lema 3.3.1. *La sucesión $\{\mathbf{Z}^n\}$ es tensa en $\mathcal{D}([0,\tau), \mathcal{M}_b^+(\mathbb{R}^+))$ si, y sólo si, para cada $\phi \in \mathcal{C}_b^+(\mathbb{R}^+)$ la sucesión asociada de procesos real-valuados $\left\{ \langle \mathbf{Z}_t^n, \phi \rangle \,\middle|\, 0 \leq t \leq \tau \right\}$ es tenso en $\mathbb{R}^+$.*

A continuación, se realizará un cálculo del cuarto momento de los incrementos de $\langle Z_t^n, \phi \rangle$. Para este propósito, se introduce la siguiente notación. Para cada $\nu \geq 1$, sea

$$d_{t,s}^{(\nu)} := \mathbf{E}\left[\left(\langle \mathbf{X}_{t+s}, \phi \rangle - \langle \mathbf{X}_t, \phi \rangle \right)^\nu \right],$$

$$k_t^{(\nu)} := (-1)^{\nu-1} \frac{\partial^\nu}{\partial \theta^\nu} K_t(\theta\phi) \mid_{\theta=0}$$

las funciones momentos

$$u_t^{(\nu)}(x) := \mathbf{E}_x\left[\langle \mathbf{X}_t, \phi\rangle^\nu\right],$$

y los momentos condicionales

$$m_{t,s}^{(\nu)}(\mathbf{X}_t) := \mathbf{E}\left[\langle \mathbf{X}_{t+s}, \phi\rangle^\nu \,\middle|\, \mathcal{F}_t\right].$$

Gracias al Lema 3.2.3 (ii), los momentos satisfacen la siguiente relación

$$u_t^{(\nu)}(x) = \frac{\partial^\nu}{\partial\theta^\nu}\{-K_t(\theta\phi)\}|_{\theta=0} = (-1)^\nu \frac{\partial^\nu}{\partial\theta^\nu}\{1 - S_t(\theta\phi)\}|_{\theta=0}.$$

Se observa que las definiciones análogas para las copias de $\mathbf{X}$ coinciden con las definidas para $\mathbf{X}$ omitiendo los superíndices. También se observa que, habiendo omitido los superíndices, $\mathbf{X}_0$ podría depender de n. Sea μ^n la distribución tal que

$$\langle \mu^n, \phi\rangle = \mathbf{E}\left[\langle \mathbf{X}_0, \phi\rangle\right] = \mathbf{E}\left[\phi\left(Y^n\right)\right].$$

Así, por ejemplo,

$$\begin{aligned}
\langle \mu_n, T_{t+s}\phi - T_t\phi\rangle &= \mathbf{E}\left[T_{t+s}\phi\left(Y^n\right) - T_t\phi\left(Y^n\right)\right] \\
&= \mathbf{E}\left[\mathbf{E}_{Y^n}\left[\langle \mathbf{X}_{t+s}, \phi\rangle\right] - \mathbf{E}_{Y^n}\left[\langle \mathbf{X}_t, \phi\rangle\right]\right] \\
&= \mathbf{E}\left[\langle \mathbf{X}_{t+s}, \phi\rangle - \langle \mathbf{X}_t, \phi\rangle\right] \\
&= d_{t,s}^{(1)}.
\end{aligned}$$

En particular, si $\mu_n = \mu_\lambda$, una distribución exponencial de parámetro λ, entonces

$$d_{t,s}^{(1)} = \langle \mu_n, T_{t+s}\phi - T_t\phi\rangle = \int_0^t \big(T_{t+s}\phi(x) - T_t\phi(x)\big)\lambda\exp\{-\lambda x\}dx = 0.$$

Lema 3.3.2. *Se satisface la siguiente relación*

$$\begin{aligned}
\mathbf{E}\left[\left|\left\langle \mathbf{Z}_{t+s}^n, \phi\right\rangle - \left\langle \mathbf{Z}_t^n, \phi\right\rangle\right|^4\right] = \frac{1}{n^3} d_{nt,ns}^4 &+ \frac{3(n-1)}{n^3}\left(d_{nt,ns}^{(2)}\right)^2 + \frac{4(n-1)}{n^3} d_{nt,ns}^{(3)} d_{nt,ns}^{(1)} \\
&+ \frac{6(n-1)(n-2)}{n^3} d_{nt,ns}^{(2)}\left(d_{nt,ns}^{(1)}\right)^2 \\
&+ \frac{(n-1)(n-2)(n-3)}{n^3}\left(d_{nt,ns}^{(1)}\right)^4. \qquad (3.3.1)
\end{aligned}$$

Demostración. Sean

$$f(k) := \sum_{j=1}^k \left(\left\langle \mathbf{X}_{t+s}^{j,k}, \phi\right\rangle - \left\langle \mathbf{X}_t^{j,k}, \phi\right\rangle\right) \quad \text{y} \quad g(k) := \left\langle \mathbf{X}_{t+s}^{k,n}, \phi\right\rangle$$

para $k = 1, 2, \ldots, n$. Como

$$\left\langle \mathbf{Z}_{t+s}^n, \phi\right\rangle - \left\langle \mathbf{Z}_t^n, \phi\right\rangle = \frac{1}{n}\sum_{j=1}^n \Big(\left\langle \mathbf{X}_{t+s}^{j,n}, \phi\right\rangle - \left\langle \mathbf{X}_{t+s}^{j,n}, \phi\right\rangle\Big),$$

entonces

$$\mathbf{E}\left[\left|\left\langle \mathbf{Z}^n_{t+s},\phi\right\rangle - \langle \mathbf{Z}^n_t,\phi\rangle\right|^4\right] = \frac{1}{n^4}\mathbf{E}\left[f(n)^4\right] = \frac{1}{n^4}\mathbf{E}\left[\big(f(n-1)+g(n)\big)^4\right].$$

Por otro lado, por la expansión binomial y la independencia de las copias, el último valor esperado satisface la siguiente fórmula recursiva

$$\begin{aligned}\mathbf{E}\left[\big(f(n-1)+g(n)\big)^4\right] &= \mathbf{E}\left[f(n-k)^4\right] + 4kd^{(1)}_{t,s}\mathbf{E}\left[f(n-k)^3\right]\\ &\quad + 6k\left((k-1)\left(d^{(1)}_{t,s}\right)^2 + d^{(2)}_{t,s}\right)\mathbf{E}\left[f(n-k)^2\right]\\ &\quad + 4k(n-1)d^{(1)}_{t,s}d^{(3)}_{t,s} + 6(k^2-k)\big(2n-(2+k)\big)\left(d^{(1)}_{t,s}\right)^2 d^{(2)}_{t,s}\\ &\quad + \left(k^3-k\right)(4n-(3k+6))\left(d^{(1)}_{t,s}\right)^4 + 3(k^2-k)\left(d^{(2)}_{t,s}\right)^2 + kd^{(4)}_{t,s}\end{aligned}$$

para $k = 1,\ldots,n-1$. En particular, para $k = n-1$, se puede escribir como sigue

$$\begin{aligned}\mathbf{E}\left[\big(f(n-1)+g(n)\big)^4\right] &= \mathbf{E}\left[f(1)^4\right] + 4(n-1)d^{(1)}_{t,s}\mathbf{E}\left[f(1)^3\right]\\ &\quad + 6(n-1)\left((n-2)\left(d^{(1)}_{t,s}\right)^2 + d^{(2)}_{t,s}\right)\mathbf{E}\left[f(1)^2\right]\\ &\quad + 4(n-1)^2 d^{(1)}_{t,s}d^{(3)}_{t,s}\\ &\quad + 6\left((n-1)^2-(n-1)\right)\big(2n-(n+1)\big)\left(d^{(1)}_{t,s}\right)^2 d^{(2)}_{t,s}\\ &\quad + \left((n-1)^3-(n-1)\right)(4n-(3(n-1)+6))\left(d^{(1)}_{t,s}\right)^4\\ &\quad + 3\left((n-1)^2-(n-1)\right)\left(d^{(2)}_{t,s}\right)^2 + (n-1)d^{(4)}_{t,s}\end{aligned}$$

Luego, agrupando términos y usando la identidad $\mathbf{E}\left[f(1)^k\right] = d^{(k)}_{t,s}$ para $k = 2,3,4$, se obtiene

$$\begin{aligned}\mathbf{E}\left[\big(f(n-1)+g(n)\big)^4\right] &= nd^{(4)}_{t,s} + 4n(n-1)d^{(1)}_{t,s}d^{(3)}_{t,s} + 6n(n-1)(n-2)\left(d^{(1)}_{t,s}\right)^2 d^{(2)}_{t,s}\\ &\quad + 3n(n-1)\left(d^{(2)}_{t,s}\right)^2.\end{aligned}$$

Por lo tanto,

$$\begin{aligned}n^4\mathbf{E}\left[\left|\left\langle \mathbf{Z}^n_{t+s},\phi\right\rangle - \langle \mathbf{Z}^n_t,\phi\rangle\right|^4\right] &= nd^{(4)}_{t,s} + 4n(n-1)d^{(1)}_{t,s}d^{(3)}_{t,s} + 6n(n-1)(n-2)\left(d^{(1)}_{t,s}\right)^2 d^{(2)}_{t,s}\\ &\quad + 3n(n-1)\left(d^{(2)}_{t,s}\right)^2,\end{aligned}$$

lo cual resulta ser (3.3.1). ■

Lema 3.3.3. *Para $\nu \geq 2$,*

(i) $u^{(\nu)}_t = T_{t-s}(\phi^\nu) + \lambda\sum_{k=1}^{\nu-1}\binom{\nu}{k}\int_0^t T_t\left(u^{(k)}_s\right)u^{(\nu-k)}_s(0)ds,$ *con* $u^{(1)}_t = T_t\phi,$

(ii) $d_{t,s}^{(\nu)} = \mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^\nu\right] + \sum_{k=2}^{\nu} (-1)^{\nu-k} \binom{\nu}{k} \mathbf{E}\left[\langle \mathbf{X}_t, \phi\rangle^{\nu-k}\left(m_{t,s}^{(k)}(\mathbf{X}_t) - \langle \mathbf{X}_t, T_s\phi\rangle^k\right)\right].$

Para $\nu \geq 1$ *y con* $m_{t,s}^0 = 1$,

(iii) $m_{t,s}^{(\nu)}(\mathbf{X}_t) = \sum_{k=0}^{\nu-1} \binom{\nu-1}{k} \left\langle \mathbf{X}_t, k_s^{(\nu-k)}\right\rangle m_{t,s}^{(k)}(\mathbf{X}_t)$, *casi en todas partes.*

Demostración. **(i)** Primero se observa que $S_t\phi$ satisface la siguiente ecuación integral

$$S_t\phi(x) = T_t\left(1 - \exp\{-\phi\}\right)(x) - \lambda \int_0^t T_{t-s}\left(S_s\phi(x)S_s\phi(0)\right)ds, \quad x \in \mathbb{R}^+. \tag{3.3.2}$$

Para verlo, sea $x \in \mathbb{R}^+$ y defínase la función f en $\mathbb{R}^+$ por

$$f(s) = 1 + \lambda \int_0^s T_r\left(1 - \exp\{-\phi\}\right)(0)dr.$$

Entonces, utilizando la propiedad de semigrupo,

$$\begin{aligned}
-\lambda \int_0^t T_{t-s}\big(S_s\phi(x)S_s\phi(0)\big)ds &= -\lambda \int_0^t T_{t-s}\big(S_s\phi(x)S_s\phi(0)\big)ds \\
&= -\lambda \int_0^t T_{t-s}\left(T_s\left(1 - \exp\{-\phi\}\right)(x)f(s)^{-1}S_s\phi(0)\right)ds \\
&= -\lambda \int_0^t T_t\left(1 - \exp\{-\phi\}\right)(x)f(s)^{-1}S_s\phi(0)ds.
\end{aligned}$$

Reagrupando términos y sustituyendo el valor de $S_s\phi(0)$, se obtiene

$$\begin{aligned}
-\lambda \int_0^t T_{t-s}\big(S_s\phi(x)S_s\phi(0)\big)ds &= -\lambda \int_0^t T_t\left(1 - \exp\{-\phi\}\right)(x)f(s)^{-1}S_s\phi(0)ds \\
&= -T_t\left(1 - \exp\{-\phi\}\right)(x) \int_0^t \lambda f(s)^{-1}S_s\phi(0)ds \\
&= -T_t\left(1 - \exp\{-\phi\}\right)(x) \int_0^t \lambda T_s\left(1 - \exp\{-\phi\}\right)(0)f(s)^{-2}ds.
\end{aligned}$$

Luego, notando que $f'(s) = \lambda T_s\left(1 - \exp\{-\phi\}\right)(0)$, un resultado clásico de primera sustitución para integrales implica

$$\begin{aligned}
-\lambda \int_0^t T_{t-s}\big(S_s\phi(x)S_s\phi(0)\big)ds &= -T_t\left(1 - \exp\{-\phi\}\right)(x) \int_0^t \lambda T_s\left(1 - \exp\{-\phi\}\right)(0)f(s)^{-2}ds \\
&= -T_t\left(1 - \exp\{-\phi\}\right)(x) \int_0^t f'(s)f(s)^{-2}ds \\
&= -T_t\left(1 - \exp\{-\phi\}\right)(x) \int_{f(0)}^{f(t)} x^{-2}dx \\
&= -T_t\left(1 - \exp\{-\phi\}\right)(x)\big[f(0)^{-1} - f(t)^{-1}\big].
\end{aligned}$$

La expresión (3.3.2) se sigue después realizar algunos despejes y sustituir los valores de $f(0)$ y $f(t)$. Ahora, utilizando la fórmula

$$\frac{\partial}{\partial\theta} T_t\left(1-\exp\{-\theta\phi\}\right)|_{\theta=0} = T_t\phi,$$

aplicando $S_t\phi$ a $\theta\phi$ y derivando ν veces respecto a θ la expresión anterior, se consigue

$$u_t^{(\nu)} = T_t(\phi^\nu) + \lambda\int_0^t T_{t-s}\left(\sum_{k=1}^{\nu-1}\binom{\nu}{k} u_s^k u_s^{\nu-k}(0)\right) ds, \quad \nu \geq 2,$$

y en consecuencia, por la linealidad del operador T, se obtiene (i).

(ii) Utilizando un argumento similar al desarrollado en la prueba del Lema 3.3.2, por la expansión binomial se tiene

$$d_{t,s}^{(\nu)} = \sum_{k=1}^{\nu}(-1)^{\nu-k}\binom{\nu}{k}\mathbf{E}\left[\langle \mathbf{X}_t, \phi\rangle^{\nu-k} m_{t,s}^{(k)}(\mathbf{X}_t)\right]. \tag{3.3.3}$$

Se recuerda que la definición de los momentos $u_t^{(\nu)}$ puede ser extendido a funciones en $\mathcal{C}_b(\mathbb{R}^+)$. Así, por ejemplo,

$$\mathbf{E}\left[\langle \mathbf{X}_t, T_t\phi - \phi\rangle^\nu\right] = \sum_{k=1}^{\nu}(-1)^{\nu-k}\binom{\nu}{k}\mathbf{E}\left[\langle \mathbf{X}_t, \phi\rangle^{\nu-k}\langle \mathbf{X}_t, T_s\phi\rangle^k\right].$$

Usando esta última expresión con (3.3.3) se obtiene

$$d_{t,s}^{(\nu)} - \mathbf{E}\left[\langle \mathbf{X}_t, T_t\phi - \phi\rangle^\nu\right] = \sum_{k=2}^{\nu}(-1)^{\nu-k}\binom{\nu}{k}\mathbf{E}\left[\langle \mathbf{X}_t, \phi\rangle^{\nu-k}\left(m_{t,s}^{(k)}(\mathbf{X}_t) - \langle \mathbf{X}_t, T_s\phi\rangle^k\right)\right],$$

y con ello (ii).

(iii) Primero, se observa que el Lema 3.2.2 (iii) justifica la siguiente relación

$$\begin{aligned}\frac{\partial^\nu}{\partial\theta^\nu}\mathbf{E}\left[\exp\{-\langle \mathbf{X}_{t+s}, \theta\phi\rangle\}\,\middle|\,\mathcal{F}_t\right] &= \frac{\partial^\nu}{\partial\theta^\nu}\exp\big\{-\langle \mathbf{X}_t, K_s(\theta\phi)\rangle\big\}\\ &= \frac{\partial^{\nu-1}}{\partial\theta^{\nu-1}}\exp\left\{-\langle \mathbf{X}_t, K_s(\theta\phi)\rangle\left\langle \mathbf{X}_t, -\frac{\partial}{\partial\theta}K_s(\theta\phi)\right\rangle\right\}.\end{aligned}$$

El resultado se sigue diferenciando el producto $\nu-1$ veces y haciendo $\theta = 0$. ■

Ahora, conviene dar algunas expresiones cerradas para $m_{t,s}^{(\nu)}(\mathbf{X}_t)$, $d_{t,s}^{(\nu)}$, $u_t^{(\nu)}$ y $k_t^{(\nu)}$. Por el Lema 3.3.3 se tiene

$$\begin{aligned}m_{t,s}^{(1)}(\mathbf{X}_t) &= \langle \mathbf{X}_t, T_s\phi\rangle,\\ m_{t,s}^{(2)}(\mathbf{X}_t) &= \left\langle \mathbf{X}_t, k_s^{(2)}\right\rangle + \left\langle \mathbf{X}_t, k_s^{(1)}\right\rangle^2,\\ m_{t,s}^{(3)}(\mathbf{X}_t) &= \left\langle \mathbf{X}_t, k_s^{(3)}\right\rangle + 3\left\langle \mathbf{X}_t, k_s^{(1)}\right\rangle\left\langle \mathbf{X}_t, k_s^{(2)}\right\rangle + \left\langle \mathbf{X}_t, k_s^{(1)}\right\rangle^3,\\ m_{t,s}^{(4)}(\mathbf{X}_t) &= \left\langle \mathbf{X}_t, k_s^{(4)}\right\rangle + 4\left\langle \mathbf{X}_t, k_s^{(1)}\right\rangle\left\langle \mathbf{X}_t, k_s^{(3)}\right\rangle + 6\left\langle \mathbf{X}_t, k_s^{(1)}\right\rangle^2\left\langle \mathbf{X}_t, k_s^{(2)}\right\rangle\\ &\quad + \left\langle \mathbf{X}_t, k_s^{(1)}\right\rangle^4 + \left\langle \mathbf{X}_t, k_s^{(2)}\right\rangle^2.\end{aligned}$$

Entonces, usando estos valores en el Lema 3.3.3 (ii),

$$\begin{aligned}
d_{t,s}^{(1)} &= \mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle\right], \\
d_{t,s}^{(2)} &= \mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^2\right] + \mathbf{E}\left[\left\langle \mathbf{X}_t, k_s^{(2)}\right\rangle\right], \\
d_{t,s}^{(3)} &= \mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^3\right] + 3\mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle \left\langle \mathbf{X}_t, k_s^{(2)}\right\rangle\right] + \mathbf{E}\left[\left\langle \mathbf{X}_t, k_s^{(3)}\right\rangle\right], \\
d_{t,s}^{(4)} &= \mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^4\right] + 6\mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^2 \left\langle \mathbf{X}_t, k_s^{(2)}\right\rangle\right] + 4\mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle \left\langle \mathbf{X}_t, k_s^{(3)}\right\rangle\right] \\
&\quad + \mathbf{E}\left[\left\langle \mathbf{X}_t, k_s^{(2)}\right\rangle^2\right] + \mathbf{E}\left[\left\langle \mathbf{X}_t, k_s^{(4)}\right\rangle\right]. \qquad (3.3.4)
\end{aligned}$$

$$\begin{aligned}
u_t^{(1)} &= T_t\phi, \\
u_t^{(2)} &= T_t\left(\phi^2\right) + 2\lambda T_t\phi \int_0^t T_s\phi(0)ds, \\
u_t^{(3)} &= T_t\left(\phi^3\right) + 3\lambda T_t\phi \int_0^t T_s\left(\phi^2\right)(0)ds + 3\lambda T_t\left(\phi^2\right)\int_0^t T_s\phi(0)ds \\
&\quad + 6\lambda^2 T_t\phi\left(\int_0^t T_s\phi(0)ds\right)^2, \\
u_t^{(4)} &= T_t\left(\phi^4\right) + 4\lambda T_t\phi \int_0^t T_s\left(\phi^3\right)(0)ds + 6\lambda T_t\left(\phi^2\right)\int_0^t T_s\left(\phi^2\right)(0)ds \\
&\quad + 4\lambda T_t\left(\phi^3\right)\int_0^t T_s\phi(0)ds + 24\lambda^2 T_t\phi\left(\int_0^t T_s\phi(0)ds\right)\left(\int_0^t T_s\left(\phi\right)^2(0)ds\right) \\
&\quad + 12\lambda^2 T_t\left(\phi^2\right)\left(\int_0^t T_s\phi(0)ds\right)^2 + 24\lambda^3 T_t\phi\left(\int_0^t T_s\phi(0)ds\right)^{3.}
\end{aligned}$$

Y ya que los primeros elementos de la sucesión $u_t^{(\nu)}$ y $k_t^{(\nu)}$ estan relacionados por

$$\begin{aligned}
k_t^{(1)} &= u_t^{(1)}, \\
k_t^{(2)} &= u_t^{(1)} - \left(u_t^{(1)}\right)^2, \\
k_t^{(3)} &= u_t^{(3)} - 3u_t^{(2)}u_t^{(1)} + 2\left(u_t^{(1)}\right)^3, \\
k_t^{(4)} &= u_t^{(4)} - 4u_t^{(3)}u_t^{(1)} - 3\left(u_t^{(2)}\right)^2 + 12u_t^{(2)}\left(u_t^{(1)}\right)^2 - 6\left(u_t^{(1)}\right)^4,
\end{aligned}$$

se tiene

$$
\begin{aligned}
k_t^{(1)} &= T_t\phi, \\
k_t^{(2)} &= T_t\left(\phi^2\right) - (T_t\phi)^2 + 2\lambda T_t\phi \int_0^t T_s\phi(0)ds, \\
k_t^{(3)} &= T_t\left(\phi^3\right) - 3T_t\phi T_t\left(\phi^2\right) + 2\left(T_t\phi\right)^3 + 3\lambda T_t\phi \int_0^t T_s\left(\phi^2\right)(0)ds \\
&\quad + 3\lambda T_t\left(\phi\right)^2 \int_0^t T_s\phi(0)ds - 6\lambda(T_t\phi)^2 \int_0^t T_s\phi(0)ds + 6\lambda^2 T_t\phi \left(\int_0^t T_s\phi(0)ds\right)^2, \\
k_t^{(4)} &= T_t\left(\phi^4\right) - 4T_t\phi T_t\left(\phi^3\right) - 3\left(T_t\left(\phi^2\right)\right)^2 + 12T_t\left(\phi^2\right)\left(T_t\phi\right)^2 - 6\left(T_t\phi\right)^4 \\
&\quad + 4\lambda T_t\phi \int_0^t T_s\left(\phi^3\right)(0)ds + 6\lambda T_t\left(\phi^2\right) \int_0^t T_s\left(\phi^2\right)(0)ds \\
&\quad + 4\lambda T_t\left(\phi^3\right) \int_0^t T_s\phi(0)ds - 24\lambda T_t\phi T_t\left(\phi^2\right) \int_0^t T_s\phi(0)ds \\
&\quad - 12\lambda\left(T_t\phi\right)^2 \int_0^t T_s\left(\phi^2\right)(0)ds + 24\lambda\left(T_t\phi^3\right) \int_0^t T_s\phi(0)ds \\
&\quad + 24\lambda^2 T_t\phi \left(\int_0^t T_s\phi(0)ds\right)\left(\int_0^t T_s\left(\phi^2\right)(0)ds\right) \\
&\quad + 12\lambda^2 T_t\left(\phi^2\right)\left(\int_0^t T_s\phi(0)ds\right)^2 - 36\lambda^2(T_t\phi)^2\left(\int_0^t T_s\phi(0)\right)^2 \\
&\quad + 24\lambda^3 T_t\phi\left(\int_0^t T_s\phi(0)ds\right)^3.
\end{aligned}
$$

Lema 3.3.4. *Sea C una constante positiva cuyo valor numerico puede cambiar y puede depender de ϕ y ϕ'. Entonces*

$$
\left|d_{t,s}^{(\nu)}\right| \leq \begin{cases} Cs, & \nu = 2, \\ C(s + s^2 + st), & \nu = 3, \\ C\left(s^3 + ts^2 + s^2(t + s + st + t^2)^{1/2} + s + s^2 + st\right), & \nu = 4. \end{cases}
$$

Demostración. Sea $\phi \in \mathcal{C}_c^{+,1}\left(\mathbb{R}^+\right)$ y C como antes. Entonces, de la expresión explícita que se tiene de $u_t^{(\nu)}$,

$$
u_t^{(\nu)} \leq \begin{cases} C\left(\|\phi\|_\infty + \int_0^t T_s\phi(0)ds\right) & \text{si} \quad \nu = 2, \\ C\left(\|\phi\|_\infty^2 + t\left(\|\phi\|_\infty + \int_0^t T_r\phi(0)dr\right)\right) & \text{si} \quad \nu = 3, \\ C\left(\|\phi\|_\infty^3 + t\left(\|\phi\|_\infty + \int_0^t T_r\phi(0)dr\right)^2\right) & \text{si} \quad \nu = 4. \end{cases}
$$

Por otro lado, se observa que $\left|\int_0^t (T_{r+s}\phi - T_r\phi)(0)dr\right|$ es menor o igual que

$$\left|\int_0^t \big(\exp\{-\lambda(r+s)\}\big(\phi(r+s)-\phi(r)\big) + \big(\exp\{-\lambda(r+s)\} - \exp\{-\lambda r\}\big)\phi(r)\big)\, dr\right|$$
$$+\left|\int_0^t \int_r^{r+s} \phi(v)\lambda \exp\{-\lambda v\} dv dr\right|,$$

que a su vez, gracias al teorema de valor medio, es menor o igual que

$$\left(\|\phi'\|_\infty + \|\phi\|_\infty\right) s + \|\phi\|_\infty s,$$

es decir,

$$\left|\int_0^t (T_{r+s}\phi - T_r\phi)(0)dr\right| \leq \left(\|\phi'\|_\infty + \|\phi\|_\infty\right) s + \|\phi\|_\infty s \leq Cs.$$

De manera análoga, se puede concluir que

$$\|T_s\phi - \phi\|_\infty \leq Cs. \tag{3.3.5}$$

Ahora, con el fin de enfatizar la función ϕ que se utiliza en la forma de la función de momentos, se escribirá $u_t[\phi](x)$ por $u_t(x)$. Con esta convención, y recordando que $\mathbf{X}_0 = \delta_Y$, se tiene

$$\mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^\nu\right] = \left\langle \mu^n, u_t^{(\nu)}\left[T_s\phi - \phi\right]\right\rangle.$$

Así, los cálculos anteriores implican

$$\left|\mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^\nu\right]\right| \leq \begin{cases} Cs, & \nu = 2, \\ C(s^2 + st), & \nu = 3, \\ C\left(s^3 + s^2 t\right), & \nu = 4. \end{cases} \tag{3.3.6}$$

Más aún, de las fórmulas explícitas obtenidas anteriormente para $k_t^{(\nu)}$, $\nu = 1, \ldots, 4$, en términos de T_t, y usando nuevamente (3.3.5), se tiene

$$\left|k_t^{(\nu)}\right| \leq \begin{cases} Ct, & \nu = 2, \\ C(t^2 + t^2), & \nu = 3, \\ C\left(t + t^2 + t^3\right), & \nu = 4, \end{cases}$$

y en consecuencia,

$$\left|\mathbf{E}\left[\left\langle \mathbf{X}_t, k_s^{(\nu)}\right\rangle\right]\right| \leq \begin{cases} Cs, & \nu = 2, \\ C(s + s^2), & \nu = 3, \\ C\left(s + s^2 + s^3\right), & \nu = 4. \end{cases} \tag{3.3.7}$$

Además, ya que $\mathbf{E}\left[\mathbf{N}_t^2\right] = (1+2\lambda t)$,

$$\left|\mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle \langle \mathbf{X}_t, k_s^{(2)}\rangle\right]\right| \le C\mathbf{E}\left[\mathbf{N}_t^2\right] s \le C(s+st), \tag{3.3.8a}$$

$$\left|\mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle \langle \mathbf{X}_t, k_s^{(3)}\rangle\right]\right| \le C\mathbf{E}\left[\mathbf{N}_t^2\right] (s+s^2) \le C\left(s+s^2+st+s^2t\right) \tag{3.3.8b}$$

y

$$\mathbf{E}\left[\langle \mathbf{X}_t, k_s^{(2)}\rangle^2\right] \le C\mathbf{E}\left[\mathbf{N}_t^{(2)}\right] s^2 \le C(s^2+s^2t). \tag{3.3.8c}$$

Finalmente, por la desiguladad de Schwarz, (3.3.6) y (3.3.8c),

$$\begin{aligned}\mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^2 \langle \mathbf{X}_t, k_s^{(2)}\rangle\right] &\le \mathbf{E}\left[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^4\right]^{1/2} \mathbf{E}\left[\langle \mathbf{X}_t, k_s^{(2)}\rangle^2\right]^{1/2} \\ &\le C(s^3+s^2t)^{1/2}(s^2+s^2t)^{1/2} \\ &= Cs^2(t+s+st+s^2)^{1/2}. \end{aligned} \tag{3.3.9}$$

Luego, usando las estimaciones (3.3.6)-(3.3.9) en (3.3.4), se obtiene el resultado.

■

Lema 3.3.5. *Sea $\phi \in \mathcal{C}_c^{+,1}(\mathbb{R}^+)$. Se asume que $\overline{\mathbf{Z}}_0 = \mu_\lambda$. Bajo las hipótesis (a) y (b) del Teorema 3.1.3 se tiene, para $0 \le t \le \tau$ y $0 \le s \le 1$,*

$$\mathbf{E}\left[\left|\langle \mathbf{Z}_{t+s}^n, \phi\rangle - \langle \mathbf{Z}_t^n, \phi\rangle\right|^4\right] \le C\left(s^2 + \frac{s}{n}\right). \tag{3.3.10}$$

Demostración. Por el Lema 3.3.4 se cumple,

$$\frac{d_{nt,ns}^{(2)}}{n} \le Cs, \qquad \left|\frac{d_{nt,ns}^{(3)}}{n^2}\right| \le C\left(s^2+st+\frac{s}{n}\right)$$

y

$$\begin{aligned}\frac{d_{nt,ns}^{(4)}}{n^3} &\le C\left(s^3+ts+s^2\left(st+t^2+\frac{s+t}{n}\right)^{1/2} + \frac{s^2+st}{n} + \frac{s}{n^2}\right) \\ &\le C\left(s^2+\frac{s}{n}\right).\end{aligned}$$

Ahora, insertando estas estimaciones en la identidad (3.3.1) del Lema 3.3.2, se obtiene

$$\mathbf{E}\left[\left|\langle \mathbf{Z}_{t+s}^n, \phi\rangle - \langle \mathbf{Z}_t^n, \phi\rangle\right|^4\right] \le C\left(s^2+\frac{s}{n}+s\left|d_{nt,ns}^{(1)}\right| + s\left(d_{nt,ns}^{(1)}\right)^2 + \left(d_{nt,ns}^{(1)}\right)^4\right). \tag{3.3.11}$$

Para finalizar, resta estimar $d_{nt,ns}^{(1)}$. Sin embargo, se tiene

$$\begin{aligned} d_{nt,ns}^{(1)} &= \int_0^\infty \left(T_{n(t+s)}\phi - T_{nt}\phi\right)(x) g_n(x)dx \\ &= \exp\{-\lambda nt\}\int_0^\infty \phi(x+nt+ns)\big(\exp\{-\lambda ns\}g_n(x) - g_n(x+ns)\big)dx \\ &\quad + \exp\{-\lambda nt\}\int_0^{ns} \phi(x+nt)\big(e(x)-g_n(x)\big)dx. \end{aligned}$$

Por consiguiente,

$$|d^{(1)}_{nt,ns}| \leq \|\phi\|_\infty \left(\int_0^\infty \big|a_n(x) - a_n(x+ns)\big| e(x)dx + \int_0^{ns} \big|g_n(x) - e(x)\big| dx \right).$$

Así, por las hipótesis (a) y (b) del Teorema 3.1.3,

$$|d^{(1)}_{nt,ns}| \leq \|\phi\|_\infty \left(\|a_n(\cdot) - a_n(\cdot + ns)\|_\infty + ns\, \|g_n - e\|_\infty \right) \leq Cs.$$

Esto, junto con (3.3.11), permite concluir la cota deseada. ■

Proposición 3.3.6. *Se asume que $\overline{\mathbf{Z}}_0$ es igual a la medida exponencial μ_λ. Entonces, bajo los supuestos (a) y (b) del Teorema 3.1.3, la sucesión $\{\mathbf{Z}^n\}$ es tensa en $\mathcal{D}([0,\tau), \mathcal{M}_b^+(\mathbb{R}^+))$.*

Demostración. Por [4], es suficiente probar que para cada ϵ y η reales positivos, existe un real positivo δ y un entero n_0 tal que

$$\mathbf{P}\left(w''\left(\langle \mathbf{Z}^n, \phi\rangle, \delta\right) \geq \epsilon\right) \leq \eta, \ n \geq n_0, \tag{3.3.12}$$

donde w'' denota el segundo módulo de continuidad definido en $[0,\tau]$(ver [4]). Aplicando el Lema 3.3.1, es suficiente verificar convergencia sólo para $\phi \in \mathcal{C}_c^{+,1}(\mathbb{R}^+) \cup \mathbf{1}$. Por el Lema 3.3.5, la cota expuesta en 3.3.10 se sostiene para $\phi \in \mathcal{C}_c^{+,1}(\mathbb{R}^+)$. Así, para $0 \leq v \leq s$ y $0 \leq t \leq \tau$,

$$\mathbf{E}\left[\left(\langle \mathbf{Z}^n_{t+s}, \phi\rangle - \langle \mathbf{Z}^n_{t+v}, \phi\rangle \right)^2 \left(\langle \mathbf{Z}^n_{t+v}, \phi\rangle - \langle \mathbf{Z}^n_{t}, \phi\rangle \right)^2\right] \leq C\left(s^2 + \frac{s}{n}\right).$$

Más aún, lo anterior permite deducir

$$\mathbf{P}\left(\left|\langle \mathbf{Z}^n_{t+s}, \phi\rangle - \langle \mathbf{Z}^n_{t+v}, \phi\rangle\right| \geq \epsilon, \left|\langle \mathbf{Z}^n_{t+v}, \phi\rangle - \langle \mathbf{Z}^n_{t}, \phi\rangle\right| \geq \epsilon\right) \leq \frac{1}{\epsilon^4} C\left(s^2 + \frac{s}{n}\right).$$

Esta última expresión, junto al Teorema 10.4 de [4], implica

$$\mathbf{P}\left(w''\left(\langle \mathbf{Z}^n, \phi\rangle, \delta\right) \geq \epsilon\right) \leq \frac{2C(\tau)}{\epsilon^4}\left(2\delta + \frac{1}{n}\right).$$

El miembro de la derecha de esta desigualdad converge a 0 cuando $n \to \infty$ y $\delta \to 0$, lo cual prueba (3.3.12). ■

Ahora, se puede completar la prueba del Teorema 3.1.3. En la Proposición 3.2.7 se establece la convergencia de todas las distribuciones finito dimensionales y en la Proposición 3.3.6, bajo los supuestos adicionales, se muestra que la sucesión de procesos aproximantes es tensa, por lo que la prueba se sigue del Teorema 2.2.20.

Bibliografía

[1] K. B. Athreya and P. E. Ney. *Branching processes*. Dover Publications, Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York; MR0373040].

[2] Krishna B. Athreya and Soumendra N. Lahiri. *Measure theory and probability theory*. Springer Texts in Statistics. Springer, New York, 2006.

[3] Robert G. Bartle. *The elements of integration and Lebesgue measure*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1995. Containing a corrected reprint of the 1966 original [ıt The elements of integration, Wiley, New York; MR0200398 (34 #293)], A Wiley-Interscience Publication.

[4] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[5] A. Bose. A law of large numbers for the scaled age distribution of linear birth-and-death processes. *Canad. J. Statist.*, 14(3):233–244, 1986.

[6] A. Bose and I. Kaj. Diffusion approximation for an age-structured population. *Ann. Appl. Probab.*, 5(1):140–157, 1995.

[7] Hans Crauel. *Random probability measures on Polish spaces*, volume 11 of *Stochastics Monographs*. Taylor & Francis, London, 2002.

[8] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[9] Luis G. Gorostiza and Jose A. López-Mimbela. The multitype measure branching process. *Adv. in Appl. Probab.*, 22(1):49–67, 1990.

[10] I. Iscoe. A weighted occupation time for a class of measure-valued branching processes. *Probab. Theory Relat. Fields*, 71(1):85–116, 1986.

[11] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.

[12] Adam Jakubowski. On the Skorokhod topology. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(3):263–285, 1986.

[13] Olav Kallenberg. *Random measures.* Akademie-Verlag, Berlin; Academic Press, Inc., London, fourth edition, 1986.

[14] Olav Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[15] David G. Kendall. Stochastic processes and population growth. *J. Roy. Statist. Soc. Ser. B.*, 11:230–264, 1949.

[16] David G. Kendall. Random fluctuations in the age-distribution of a population whose development is controlled by the simple "birth-and-death" process. *J. Roy. Statist. Soc. Ser. B.*, 12:278–285, 1950.

[17] Achim Klenke. *Probability theory.* Universitext. Springer-Verlag London, Ltd., London, 2008. A comprehensive course, Translated from the 2006 German original.

[18] Andreas E. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications.* Universitext. Springer-Verlag, Berlin, 2006.

[19] Jean-François Le Gall. *Spatial branching processes, random snakes and partial differential equations.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1999.

[20] Jean-François Le Gall. *Mouvement brownien, martingales et calcul stochastique.* Springer-Verlag Berlin Heidelberg, 2013.

[21] Zenghu Li. *Measure-valued branching Markov processes.* Probability and its Applications (New York). Springer, Heidelberg, 2011.

[22] David Pollard. *Convergence of stochastic processes.* Springer Series in Statistics. Springer-Verlag, New York, 1984.

[23] A. V. Skorohod. Limit theorems for stochastic processes. *Teor. Veroyatnost. i Primenen.*, 1:289–319, 1956.

[24] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin-New York, 1979.

[25] Ward Whitt. *Stochastic-process limits.* Springer Series in Operations Research. Springer-Verlag, New York, 2002. An introduction to stochastic-process limits and their application to queues.

Índice alfabético