

# Árboles de Galton–Watson, convergencia y genealogía de la muestra

AUTOR:

inm

PROFESOR:

Juan Carlos Pardo Millán

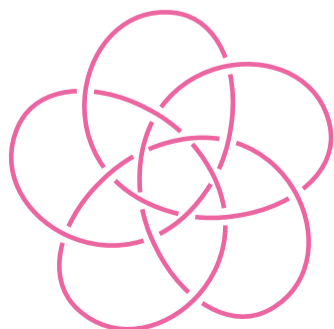
CIMAT

Guanajuato

2 de diciembre de 2023

## RESUMEN

Estas son las notas del curso *Árboles de Galton–Watson, convergencia y genealogía de la muestra*, impartido en el semestre enero–julio de 2023. En estas notas se presenta una introducción al estudio de los árboles aleatorios, haciendo énfasis en la conexión existente con teoría de excursiones.



CIMAT

---

Centro de Investigación  
en Matemáticas, A.C.

# Índice general

---

<b>Sobre la enumeración</b>	<b>4</b>
<b>I. Árboles aleatorios</b>	<b>5</b>
<b>1. Codificación de árboles aleatorios</b>	<b>6</b>
1.1. Árboles discretos . . . . .	6
1.2. Árboles de Galton–Watson . . . . .	11
1.3. Convergencia al movimiento browniano . . . . .	14
<b>2. Tiempos locales y excursiones</b>	<b>23</b>
2.1. Procesos de Markov . . . . .	23
2.2. Tiempo local . . . . .	26
2.3. Tiempo local inverso . . . . .	36
2.4. Proceso de excursiones . . . . .	39
2.A. Puntos de retención y puntos irregulares . . . . .	43
<b>3. Procesos de Lévy de variación finita</b>	<b>45</b>
3.1. Supuestos y preliminares . . . . .	45
3.2. Representación como sistema de colas . . . . .	48
3.3. Estructura de ramificación . . . . .	50
3.4. El proceso $\rho$ . . . . .	53
3.5. Propiedades de dualidad . . . . .	57
3.6. Combinando ramificación y movimiento espacial . . . . .	60
<b>4. Convergencia de funcionales de árboles de Galton–Watson</b>	<b>62</b>
4.1. Primeras aplicaciones del teorema 1.3.2 . . . . .	62
4.2. Árboles de Galton–Watson con progenie total fija . . . . .	67
4.3. Convergencia de las funciones de contorno . . . . .	71
4.A. Demostración del lema IV.2.2 . . . . .	73
<b>5. Árboles reales y su codificación mediante excursiones brownianas</b>	<b>81</b>
5.1. Árboles reales . . . . .	81
5.2. Codificación de árboles reales . . . . .	84
5.3. El árbol aleatorio continuo (CRT) . . . . .	90
5.4. La medida de excursiones de Itô . . . . .	91
5.5. Distribuciones finito-dimensionales bajo la medida de Itô . . . . .	94
5.6. Distribuciones finito-dimensionales del CRT . . . . .	100

<b>II. Genealogía de una muestra</b>	<b>104</b>
<b>6. Introducción</b>	<b>105</b>
6.1. El modelo base . . . . .	105
6.2. Resultados auxiliares . . . . .	106
<b>7. Espinas y cambios de medidas</b>	<b>107</b>
7.1. La medida de $k$ -espinas . . . . .	107
7.2. Cambio de medida . . . . .	108
7.3. Descripción de $\mathbb{Q}^{k,\top}$ . . . . .	110
7.4. Comportamiento de $\mathbb{Q}^{k,\top}$ para tiempos grandes . . . . .	115
7.5. Demostración de la proposición VII.2.2 . . . . .	116
7.6. Fórmula de Campbell . . . . .	116
<b>8. Procesos de nacimiento y muerte</b>	<b>117</b>
<b>9. Límite de escala en el caso casi-crítico</b>	<b>118</b>
<b>10. Discusión de los resultados</b>	<b>119</b>
<b>III. Apéndices</b>	<b>120</b>
<b>A. Cálculo estocástico</b>	<b>121</b>
A.1. Tiempo local para semimartingalas . . . . .	121
A.2. Tiempo local para el movimiento browniano . . . . .	127
<b>Referencias</b>	<b>129</b>

# Sobre la enumeración

---

Usamos un sistema de numeración jerárquico para ecuaciones, enunciados y figuras. La  $k$ -ésima ecuación de la sección  $j$  del capítulo  $i$  se etiqueta como  $(k)$  en el lugar donde aparece, y es citada como  $(k)$  dentro de la misma sección, como  $(j.k)$  dentro del capítulo  $i$  pero fuera de la sección  $j$  y como  $(i.j.k)$  fuera del capítulo  $i$ . El  $k$ -ésimo enunciado dentro de una sección se etiqueta como  $k$ . Enunciado en el lugar donde aparece, se cita como enunciado  $k$  dentro de la misma sección, como enunciado  $j.k$  dentro del capítulo y como enunciado  $i.j.k$  fuera del capítulo. La etiqueta de enunciado comprende los teoremas, lemas, definiciones, observaciones, proposiciones y corolarios. Esta misma notación se usa para figuras.

# Parte I.

## Árboles aleatorios

El objetivo de esta parte es construir el objeto conocido como *Continuum Random Tree*, lo cual se hace siguiendo el artículo de revisión «Random Trees and Applications» de Le Gall (2005). Para ello también se estudia el concepto de tiempos locales y excursiones de procesos de Markov siguiendo las ideas del libro *Lévy Processes* de Bertoin (1996). Asimismo, para ejemplificar el uso de las excursiones y los tiempos locales, se estudia el caso de variación finita tratado en el artículo «Branching Processes in Lévy Processes» de Le Gall y Le Jan (1998).

# Codificación de árboles aleatorios

# 1

En este capítulo se aborda la codificación de árboles aleatorios discretos a través de ciertas funciones de este. Asimismo se prueba que el reescalamiento de una de estas funciones, la cual se asocia a un bosque de árboles aleatorios de Galton–Watson independientes, converge débilmente a un movimiento browniano reflejado. La exposición se basa en las secciones 1.1 a 1.3 del artículo de revisión «Random Trees and Applications» de Le Gall (2005).

## 1.1. Árboles discretos

A lo largo del presente apartado estaremos interesados en árboles finitos, enraizados y ordenados. Estos resultan ser sub-árboles enraizados del *árbol de Ulam–Harris*, el cual es el conjunto

$$\mathcal{U} := \bigcup_{n \geq 0} \mathbb{N}^n, \quad \text{donde } \mathbb{N}^0 = \{\emptyset\},$$

equipado con el *orden genealógico*:  $u \leq v$  si y sólo si  $u \in \mathbb{N}^n$ ,  $v \in \mathbb{N}^m$ ,  $n \leq m$  y  $u_i = v_i$  para  $i \in \llbracket 1, n \rrbracket$ . Bajo el orden dado se tiene que  $\emptyset \leq u$  para cualquier  $u \in \mathcal{U}$ , motivo por el cual se conoce a  $\emptyset$  como la *raíz*. Antes de definir formalmente los objetos de interés, conviene describir ciertos elementos asociados a los árboles:

- La *longitud generación* de  $u \in \mathcal{U}$  se denota por  $|u|$  y es igual al entero  $n \geq 0$  tal que  $u \in \mathbb{N}^n$ .
- Dados  $u = (u_1, \dots, u_n)$ ,  $v = (v_1, \dots, v_m) \in \mathcal{U}$ , la *concatenación* de  $u$  y  $v$  se escribe como  $uv = (u_1, \dots, u_n, v_1, \dots, v_m)$ . En particular  $u\emptyset = \emptyset u = u$ .
- Definimos el mapeo  $\pi : \mathcal{U} \setminus \{\emptyset\} \rightarrow \mathcal{U}$  mediante  $\pi(u) = (u_1, \dots, u_{n-1})$  para  $u \in \mathbb{N}^n$ . (Intuitivamente puede pensarse en  $\pi(u)$  como el “progenitor” de  $u$ .)

El orden dado por  $\leq$  es un orden parcial, el cual puede extenderse a un orden total considerando el orden lexicográfico  $\leq$  sobre  $\mathcal{U}$ . Tener esta idea en mente será de utilidad más adelante para introducir ciertas funciones que codificarán a nuestro objeto de interés, mismo que describimos de manera concreta a continuación.

### 1. Definición

Un *árbol finito, enraizado y ordenado*  $\mathbf{t}$  es un subconjunto finito de  $\mathcal{U}$  que satisface las propiedades:

- (i)  $\emptyset \in \mathbf{t}$ .
- (ii) Para todo  $u \in \mathbf{t} \setminus \{\emptyset\}$  se cumple  $\pi(u) \in \mathbf{t}$ .
- (iii) Para cada elemento  $u \in \mathbf{t}$  existe un entero  $k_u(\mathbf{t})$  tal que  $uj \in \mathbf{t}$  siempre que  $j \in \llbracket 1, k_u(\mathbf{t}) \rrbracket$ .

De ahora en adelante se dejará de escribir el adjetivo finito salvo riesgo de confusión. Al número  $k_u(\mathbf{t})$  se le interpreta como el “número de hijos” de  $u$  en  $\mathbf{t}$ , y

diremos que  $u$  es una hoja en caso de que  $k_u(t) = 0$ . Por las palabras usadas, entendemos a cada vértice en  $t$  como un individuo de una población cuya genealogía queda descrita por  $t$ .

Al conjunto de todos los árboles enraizados y ordenados lo denotaremos por  $\mathbf{T}$ . Para cada elemento  $t \in \mathbf{T}$  definimos a la cardinalidad  $\#t$  de  $t$  como la progenie total. Procedemos a introducir dos funciones que permiten codificar un árbol, para lo cual consideraremos los elementos  $u_0 = \emptyset, \dots, u_{\#t-1}$  de un árbol  $t$  ordenados de manera lexicográfica.

La primera función que describiremos es la *función de altura*  $h_t : [0, \#t - 1] \rightarrow \mathbb{R}_+$  definida por

$$h_t(n) = |n|.$$

De esta manera, la función de altura es la sucesión de las generaciones de los individuos en  $t$  cuando los individuos se ordenan de manera lexicográfica. La segunda función en cuestión es la *función de contorno*  $C_t : [0, \zeta(t)] \rightarrow \mathbb{R}_+$ , donde  $\zeta(t) = 2(\#t - 1)$ . Para su descripción consideremos las sucesiones  $\{l_i : i \in [0, \#t - 1]\}$  y  $\{b_i : i \in [0, \zeta(t)]\}$  dadas como sigue:

- $l_0 = b_0 = 0$ ,
- para cada  $i \in [1, \#t - 1]$ ,

$$(1) \quad l_i = \sum_{j=1}^i (|u_{j-1}| - |u_j| + 2) = 2i - 2|u_i|,$$

$$b_{l_i} = 1 \text{ y}$$

$$(2) \quad b_j = -1 \quad \text{si } j \in \bigcup_{i=1}^{\#t-1} [l_{i-1} + 1, l_i] \cup [l_{\#t-1} + 1, \zeta(t)].$$

Dado lo anterior, la función de contorno queda dada por

$$C_t(n) = \sum_{i=0}^n b_i.$$

Es posible comprobar que  $C_t$  describe un camino de Dyck de longitud  $\zeta(t)$ . En ambos casos podemos considerar extensiones a “tiempo continuo” interpolando linealmente los valores entre  $n$  y  $n + 1$ , y esto se hará para la función de contorno.

A continuación realizaremos un ejemplo para que se entienda cómo es que las funciones descritas codifican a los árboles. En la figura 1 se muestran un árbol junto y sus funciones de altura y contorno asociadas. En el árbol  $t$  se muestra una sucesión de flechas, la cual nos muestra el orden en el cual se irán recorriendo los individuos en el árbol. Para construir la función de altura simplemente consideraremos la generación de cada nuevo individuo visitado en el orden indicado por la flechas. Mientras tanto, la función de contorno se construye comenzando en cero e incrementando una unidad por cada flecha hacia arriba y decreciendo una unidad por cada flecha hacia abajo. Intuitivamente, esto describe el movimiento de una partícula sobre el contorno del árbol  $t$ . Si, en cambio, comenzamos con la función de altura para construir el árbol basta fijarnos entre dos tiempos  $n_1 < n_2$  para los cuales  $h_t(n_1) = h_t(n_2)$  pues los puntos que quedan entre ellos representarán la descendencia de  $u_{n_1}$ ; cada punto con  $h_t(n_1) + 1$  será un hijo de  $u_{n_1}$ . Mientras tanto, al tener la función de contorno, cada valle representa un cambio de rama, lo cual se ve claro en el ejemplo.



Demostración: Pongamos  $j_0 = 1$  y consideremos

$$j_i = \inf \left\{ j \geq j_{i-1} + 1 : \sum_{k=j_{i-1}+1}^j m_k = j - j_{i-1} - 1 \right\}.$$

Nos enfocaremos en verificar que  $j_1$  existe y satisface las condiciones establecidas en el enunciado, pues entonces el resultado (salvo que  $j_{m_1} = p$ ) se seguirá de un argumento recursivo. Notemos que si  $m_2 = 0$ , entonces  $j_1 = 2$  y  $\{m_2\} \in \mathcal{S}$  trivialmente. Luego, veremos el caso  $m_2 > 0$ , para lo cual reescribimos

$$j_1 = \inf \left\{ j \geq 2 : \sum_{k=2}^j (m_k - 1) = -1 \right\}.$$

Ahora bien, observemos que  $(m_2 - 1) \geq 0$  y por hipótesis tenemos

$$\sum_{k=2}^p (m_k - 1) = -m_1 \leq -1.$$

Debido a que  $(m_k - 1) \geq -1$ , debe ser el caso que exista un índice  $j \in \llbracket 3, p \rrbracket$  tal que  $\sum_{k=2}^j (m_k - 1) = -1$ . Por ende  $j_1$  está bien definido y para  $k \in \llbracket 2, j_1 - 1 \rrbracket$ ,  $\sum_{i=2}^k (m_i - 1) \geq 0$ , o de forma equivalente  $\sum_{i=1}^{k-1} m_{i+1} \geq k - 1$ . Asimismo se verifica que  $\sum_{i=1}^{j_1-1} m_{i+1} = j_1 - 2$ , de modo que  $(m_2, \dots, m_{j_2}) \in \mathcal{S}$ . Por otra parte, debe ser el caso que  $j_1 \leq p - m_1 + 1$  pues en caso contrario, es decir si  $j_1 > p - m_1 + 1$ , tendríamos que

$$\sum_{i=1}^p m_i \geq \sum_{i=1}^{j_1} m_i = m_1 + j_1 - 2 > p - 1,$$

lo cual viola la hipótesis  $\{m_i : i \in \llbracket 1, p \rrbracket\} \in \mathcal{S}$ .

Para concluir resta verificar que  $j_{m_1} = p$  pues de los argumentos previos únicamente sabemos que  $j_{m_1} \leq p$ . Ahora, si fuese el caso que  $j_{m_1} < p$  nuevamente llegaríamos a una contradicción, pues por hipótesis  $\sum_{i=1}^{j_{m_1}} m_i \geq j_{m_1}$  pero por otra parte

$$\sum_{i=1}^{j_{m_1}} m_i = m_1 + \sum_{i=1}^{m_1} \sum_{k=j_{i-1}+1}^{j_i} m_k = m_1 + \sum_{i=1}^{m_1} (j_i - j_{i-1} - 1) = j_{m_1} - j_0 - j_{m_1} - 1.$$

Así,  $j_{m_1} = p$  debe suceder, lo cual nos da el resultado deseado. ■

### 3. Proposición

*El mapeo*

$$\Phi : \mathbf{t} \mapsto (k_{u_0}(\mathbf{t}), \dots, k_{u_{\#t-1}}(\mathbf{t}))$$

*define una biyección entre  $\mathbf{T}$  y  $\mathcal{S}$ .*

Demostración: Debido a que  $k_u(\mathbf{t})$  cuenta el número de hijos que tiene  $u$  en  $\mathbf{t}$ , se sigue que  $\sum_{i=0}^{\#t-1} k_{u_i}(\mathbf{t}) = \#t - 1$  pues  $\emptyset$  no se cuenta como hijo. En adición, para  $i \in \llbracket 0, \#t - 2 \rrbracket$ ,  $\sum_{l=0}^i k_{u_l}(\mathbf{t})$  cuenta el número de hijos de  $u_0, \dots, u_i$ , entre los cuales se encuentran  $u_1, \dots, u_{i+1}$ , de modo que

$$\sum_{l=0}^i k_{u_l}(\mathbf{t}) \geq i + 1.$$

Así,  $\Phi(\mathbf{T}) \subset \mathcal{S}$ .

Verifiquemos que  $\Phi$  es biyectiva, comenzando con la inyectividad. Dados dos árboles  $\mathbf{t}_1 \neq \mathbf{t}_2$  podemos suponer sin pérdida de generalidad que  $\mathbf{t}_1 \setminus \mathbf{t}_2 \neq \emptyset$ . Al ordenar a los individuos  $u_0, \dots, u_{\#\mathbf{t}_1-1}$  en  $\mathbf{t}_1$  de manera lexicográfica, consideremos el índice del primer individuo en  $\mathbf{t}_1$  que no se encuentre en  $\mathbf{t}_2$ ,  $\hat{i} = \min\{i \in \llbracket 1, \#\mathbf{t}_1 - 1 \rrbracket : u_i \notin \mathbf{t}_2\}$ . Sea  $\hat{j} < \hat{i}$  tal que  $u_{\hat{j}} = \pi(u_{\hat{i}})$ , entonces  $k_{u_{\hat{j}}}(\mathbf{t}_2) < k_{u_{\hat{j}}}(\mathbf{t}_1)$  puesto que en caso contrario existiría  $l \in \llbracket 1, k_{u_{\hat{j}}}(\mathbf{t}_2) \rrbracket$  con la propiedad  $u_{\hat{i}} = u_{\hat{j}}l \in \mathbf{t}_2$ . Por ende,  $\Phi(\mathbf{t}_1) \neq \Phi(\mathbf{t}_2)$  cuando  $\mathbf{t}_1 \neq \mathbf{t}_2$ .

Para probar que  $\Phi$  es sobreyectiva procederemos por inducción (fuerte) sobre la longitud de una sucesión  $\tilde{m} = (m_1, \dots, m_p) \in \mathcal{S}$ . La única sucesión de longitud uno en  $\mathcal{S}$  es  $(0)$ , para la cual se tiene que  $\Phi(\{\emptyset\}) = (0)$ , dándonos el caso base. Para el paso inductivo tomemos  $(m_1, \dots, m_p) \in \mathcal{S}$  con  $m_1 \geq 1$ . En vista del lema 2, existen índices  $0 = j_1 < \dots < j_{m_1} = p$  tales que  $\tilde{m}_i = (m_{j_{i-1}+1}, \dots, m_{j_i}) \in \mathcal{S}$ . Así, por la hipótesis de inducción, existen  $\mathbf{t}_1, \dots, \mathbf{t}_{m_1} \in \mathbf{T}$  tales que  $\Phi(\mathbf{t}_i) = \tilde{m}_i$ . Finalmente consideremos

$$\mathbf{t} = \{\emptyset\} \cup \bigcup_{i=1}^{m_1} i\mathbf{t}_i, \quad \text{donde } i\mathbf{t}_i = \{iu : u \in \mathbf{t}_i\}.$$

Por construcción  $\Phi(\mathbf{t}) = \tilde{m}$ , y en consecuencia  $\Phi(\mathbf{T}) = \mathcal{S}$ . ■

Lo que la proposición 3 nos dice es que podemos codificar un árbol  $\mathbf{t} \in \mathbf{T}$  a través del número de hijos que tienen los individuos cuando los individuos se ordenan de manera lexicográfica, dando un elemento  $\Phi(\mathbf{t}) = (m_1, \dots, m_p) \in \mathcal{S}$ . En adelante también consideraremos las sucesiones de enteros

$$x_n = \sum_{i=1}^n (m_i - 1), \quad n \in \llbracket 0, p \rrbracket,$$

que ya habían aparecido en la demostración del lema 2, y satisfacen las propiedades

- $x_0 = 0$  y  $x_p = -1$ ,
- $x_n \geq 0$  para cada  $n \in \llbracket 0, p-1 \rrbracket$  y
- $x_n - x_{n-1} \geq -1$  para cada  $n \in \llbracket 1, p \rrbracket$ .

A este tipo de sucesiones de la conoce como *camino de Łukasiewicz*. Es sencillo verificar que hay una biyección entre  $\mathcal{S}$  y los caminos de Łukasiewicz, por lo cual la proposición 3 implica la existencia de una biyección entre  $\mathbf{T}$  y caminos de Łukasiewicz. Usando esta biyección, resulta que existe una relación entre la función de altura  $h_{\mathbf{t}}$  y el camino de Łukasiewicz generado por  $\mathbf{t}$ .

#### 4. Proposición

La función de altura  $h_{\mathbf{t}}$  de un árbol  $\mathbf{t}$  se relaciona con el camino de Łukasiewicz de  $\mathbf{t}$  por medio de la fórmula

$$h_{\mathbf{t}}(n) = \#\left\{j \in \llbracket 0, n-1 \rrbracket : x_j = \inf_{j \leq l \leq n} x_l\right\},$$

donde  $n \in \llbracket 0, \#\mathbf{t} - 1 \rrbracket$ .

Demostración: Observemos que para  $u, v \in \mathbf{t}$ ,  $u < v$  si y sólo si existe  $j \in \llbracket 1, |v| \rrbracket$  con la propiedad  $u = \pi^j(v)$ , de modo que  $v$  tiene  $|v|$  ancestros. Luego,

$$h_{\mathbf{t}}(n) = \#\left\{j \in \llbracket 0, n-1 \rrbracket : u_j < u_n\right\}.$$

En consecuencia, basta probar que para  $j \in \llbracket 0, n-1 \rrbracket$  la condición  $u_j < u_n$  es equivalente a  $x_j = \inf_{j \leq l \leq n} x_l$  o bien  $x_l - x_j \geq 0$  para toda  $l \in \llbracket j, n \rrbracket$ .

Recordando la prueba del lema 2 y cómo se aplicó en la demostración de la proposición 3, se deduce que el primer índice  $k$  para el cual  $\sum_{i=j}^k (m_i - 1) = -1$  nos dice que  $u_{k-1}$  es el último descendiente de  $u_{j-1}$  o, de forma equivalente, que  $u_k$  es el primer individuo que no es descendiente de  $u_{j-1}$ . Consecuentemente, que

$$0 \leq x_l - x_j = \sum_{i=j+1}^l (m_i - 1) \quad \text{para toda } l \in \llbracket j, n \rrbracket,$$

es equivalente a que  $u_j \leq u_l$  para toda  $l \in \llbracket j, n \rrbracket$ , lo cual sucede si, solamente si,  $u_j < u_n$ . ■

## 1.2. Árboles de Galton–Watson

Habiendo entendido la relación entre elementos de  $\mathbf{T}$  con ciertas funciones que permiten codificarlos, las usaremos para estudiar árboles aleatorios. Para definir estos objetos consideraremos una distribución de reproducción  $\mu$  crítica o subcrítica; esto es, una medida de probabilidad sobre  $\mathbb{N}_0$  con

$$\sum_{k=0}^{\infty} k\mu(k) \leq 1,$$

excluyendo el caso trivial donde  $\mu(1) = 1$ .

Para la construcción de árboles de Galton–Watson, consideraremos una colección  $\{K_u : u \in \mathcal{U}\}$  de variables aleatorias independientes con distribución común  $\mu$ . El árbol de Galton–Watson con distribución de reproducción  $\mu$  (GWT( $\mu$ )) al subconjunto aleatorio  $\theta$  de  $\mathcal{U}$  definido por

$$\theta = \left\{ u = u^1 \cdots u^n \in \mathcal{U} : u^j \leq K_{u^1 \dots u^{j-1}} \text{ para cada } j \in \llbracket 1, n \rrbracket \right\}.$$

### 1. Proposición

Con probabilidad uno  $\theta$  es un árbol. En adición, si para cada  $n \geq 0$  definimos

$$Z_n = \#\{u \in \theta : |u| = n\},$$

el proceso  $\{Z_n : n \geq 0\}$  es un proceso de Galton–Watson con distribución de reproducción  $\mu$  y valor inicial  $Z_0 = 1$ .

Demostración: Que  $\theta$  es un árbol con probabilidad uno se deduce de su definición. Por otra parte, como  $\{u \in \theta : |u| = 0\} = \{\emptyset\}$ ,  $Z_0 = 1$ . Más aún, observando que  $K_u = k_u(\theta)$  para cada  $u \in \theta$  y que  $\{K_u : u \in \mathcal{U}, |u| = n\}$  es independiente de la colección  $\{K_u : u \in \mathcal{U}, |u| \leq n-1\}$ , para  $f : \mathbb{Z}_+ \rightarrow \mathbb{R}$  acotada tendremos que

$$\begin{aligned} \mathbb{E}_1[f(Z_{n+1})\mathbb{1}_{\{Z_1 \in A_1, \dots, Z_n = l_n\}}] &= \sum_{l_1 \in A_1} \cdots \sum_{l_{n-1} \in A_{n-1}} \mathbb{E}_1 \left[ f \left( \sum_{u \in \theta, |u|=n} K_u \right) \mathbb{1}_{\{Z_1 = l_1, \dots, Z_n = l_n\}} \right] \\ &= \sum_{l_1 \in A_1} \cdots \sum_{l_{n-1} \in A_{n-1}} \mathbb{E}_1 \left[ \int_{\mathbb{N}_0} f(x) d\mu^{*l_n}(x) \mathbb{1}_{\{Z_1 = l_1, \dots, Z_n = l_n\}} \right] \\ &= \mathbb{P}_1(Z_1 \in A_1, \dots, Z_{n-1} \in A_{n-1}, Z_n = l_n) \mathbb{E}_n[f(Z_1)]. \end{aligned}$$

Por lo tanto, se sigue que  $\{Z_n : n \geq 0\}$  es una cadena de Markov homogénea con probabilidades de transición dadas por  $p_{nm} = \mu^{*n}(m)$ , lo cual corresponde al proceso de Galton–Watson requerido. ■

De la proposición 1 se sigue que bajo el supuesto de que la distribución de reproducción  $\mu$  sea crítica o subcrítica implica la finitud del árbol  $\theta$ , pues con probabilidad uno,  $Z_n = 0$  para  $n$  suficientemente grande.

Para describir la distribución  $\Pi_\mu$  de  $\theta$  en  $\mathbf{T}$ , haremos uso de que cada individuo en  $\theta$  genera un nuevo árbol aleatorio. En pos a describir formalmente lo que sucede, dados un árbol  $\mathbf{t}$  y  $j \in \llbracket 1, k_\emptyset(\mathbf{t}) \rrbracket$  denotamos por  $T_j \mathbf{t}$  al árbol trasladado en  $j$ , es decir

$$T_j \mathbf{t} = \{u \in \mathcal{U} : ju \in \mathbf{t}\}.$$

Es claro que  $T_j \mathbf{t}$  es un árbol. Dada esta notación, podemos caracterizar  $\Pi_\mu$  mediante las propiedades:

- $\Pi_\mu(k_\emptyset = j) = \mu(j)$  para  $j \in \mathbb{N}_0$ .
- Para cada  $j \in \mathbb{N}$  con  $\mu(j) > 0$ , los árboles  $T_1 \mathbf{t}, \dots, T_j \mathbf{t}$  son independientes bajo la medida de probabilidad  $\Pi_\mu(d\mathbf{t} \mid k_\emptyset = j)$ , con distribución condicional  $\Pi_\mu$ .

La propiedad (ii) se puede entender como que cada descendiente de la raíz  $\emptyset$  genera un nuevo árbol de Galton–Watson, el cual será independiente de los árboles generados por los demás descendientes de  $\emptyset$ . Sin usar esta caracterización, es posible dar una fórmula explícita para la distribución  $\Pi_\mu$ .

## 2. Proposición

Para todo  $\mathbf{t} \in \mathbf{T}$ ,

$$\Pi_\mu(\mathbf{t}) = \prod_{u \in \mathbf{t}} \mu(k_u(\mathbf{t})).$$

Demostración: En vista de la caracterización de un árbol  $\mathbf{t}$  por medio de la cantidad de descendientes que tiene cada individuo en  $\mathbf{t}$ , se deduce que

$$\{\theta = \mathbf{t}\} = \{K_u = k_u(\mathbf{t}) \text{ para cada } u \in \mathbf{t}\} = \bigcap_{u \in \mathbf{t}} \{K_u = k_u(\mathbf{t})\}.$$

Así, usando la independencia de los elementos en  $\{K_u : u \in \mathcal{U}\}$  obtenemos

$$\Pi_\mu(\mathbf{t}) = \mathbb{P}(\theta = \mathbf{t}) = \prod_{u \in \mathbf{t}} \mathbb{P}(K_u = k_u(\mathbf{t})) = \prod_{u \in \mathbf{t}} \mu(k_u(\mathbf{t})). \quad \blacksquare$$

## 3. Proposición

Sea  $\theta$  un GWT( $\mu$ ). Entonces

$$\Phi(\theta) \stackrel{d}{=} (M_1, \dots, M_T),$$

donde  $\{M_n : n \geq 1\}$  es una colección de variables aleatorias independientes con distribución común  $\mu$ .

$$T = \inf\{n \geq 1 : M_1 + \dots + M_n < n\}.$$

Demostración: Tomemos  $\theta$  un GWT( $\mu$ ) y denotemos por  $U_0 = \emptyset, U_1, \dots, U_{\#\theta-1}$  a sus elementos ordenados de forma lexicográfica. Entonces

$$\Phi(\theta) = (K_{U_0}, \dots, K_{U_{\#\theta-1}}).$$

De la proposición 1.3 sabemos que  $K_{U_0} + \cdots + K_{U_n} \geq n + 1$  para toda  $n \in \llbracket 0, \#\theta - 2 \rrbracket$  y  $K_{U_0} + \cdots + K_{U_{\#\theta-1}} = \#\theta - 1$ .

La idea es demostrar que  $K_{U_0}, \dots, K_{U_{\#\theta-1}}$  conforman una colección de variables aleatorias independientes con distribución  $\mu$ . El problema es que  $\#\theta - 1$  es aleatorio de modo que no se puede usar el hecho de que  $\{K_u : u \in \mathcal{U}\}$  son independientes e idénticamente distribuidas. Para darle la vuelta al problema definiremos

$$U_p = U_{\#\theta-1} \underbrace{1 \cdots 1}_{p-\#\theta+1} \quad \text{para } p \geq \#\theta,$$

usando esta extensión para probar que  $K_{U_0}, \dots, K_{U_p}$  son independientes con distribución  $\mu$  para cada  $p \geq 0$ .

Procedemos por inducción sobre  $p$ , observando que  $U_0 = \emptyset$  y  $U_1 = 1$  son deterministas y por ende nos dicen que  $K_{U_0}, \dots, K_{U_p}$  son aleatorios. Luego, para  $p \geq 2$  supongamos que el resultado deseado es cierto para  $p - 1$ . Debido a la construcción de  $\theta$ , es decir que  $u \in \theta$  si y sólo si  $u^{|u|-j+1} \leq K_{\pi^j(u)}$  para  $j \in \llbracket 1, |u| \rrbracket$ , se sigue que para  $u \in \mathcal{U}$  dado, es posible escribir al conjunto

$$\{U_p = u\} \cap \{\#\theta > p\}$$

en términos de  $\{K_v : v < u\}$ , de modo que es  $\sigma(K_v, v < u)$ -medible. De manera similar se verifica que

$$\{U_p = u\} \cap \{\#\theta \leq p\} \in \sigma(K_v : v < u).$$

En consecuencia  $\{U_p = u\}$  es medible con respecto a  $\sigma(K_v, v < u)$ .

Usando lo anterior y la independencia de las variables aleatorias  $\{K_v : v \in \mathcal{U}\}$ , al tomar  $g_0, \dots, g_p : \mathbb{N}_0 \rightarrow \mathbb{R}_+$  medibles obtenemos

$$\begin{aligned} & \mathbb{E}[g(K_{U_0}) \cdots g_p(K_{U_p})] \\ &= \sum_{u_0 < \cdots < u_p} \mathbb{E}[\mathbb{1}_{\{U_0=u_0, \dots, U_p=u_p\}} g_0(K_{u_0}) \cdots g_{p-1}(K_{u_{p-1}}) \mathbb{E}[g_p(K_{u_p}) \mid K_v, v < u_p]] \\ &= \sum_{u_0 < \cdots < u_p} \mathbb{E}[\mathbb{1}_{\{U_0=u_0, \dots, U_p=u_p\}} g_0(K_{u_0}) \cdots g_{p-1}(K_{u_{p-1}})] \mathbb{E}[g_p(K_{u_p})]. \end{aligned}$$

Puesto que  $\mathbb{E}[g_p(K_{u_p})] = \int_{\mathbb{N}_0} g_p(x) d\mu(x) = \mu(g_p)$  no depende de  $u_p$ , hemos demostrado que

$$\mathbb{E}[g_0(K_{U_0}) \cdots g_p(K_{U_p})] = \mu(g_p) \sum_{u_0 < \cdots < u_p} \mathbb{E}[\mathbb{1}_{\{U_0=u_0, \dots, U_p=u_p\}} g_0(K_{u_0}) \cdots g_{p-1}(K_{u_{p-1}})],$$

y en particular, tomando  $g_p = 1$ , que

$$\mathbb{E}[g_0(K_{U_0}) \cdots g_{p-1}(K_{U_{p-1}})] = \sum_{u_0 < \cdots < u_p} \mathbb{E}[\mathbb{1}_{\{U_0=u_0, \dots, U_p=u_p\}} g_0(K_{u_0}) \cdots g_{p-1}(K_{u_{p-1}})].$$

En consecuencia

$$\mathbb{E}\left[\prod_{i=0}^p g_i(K_{U_i})\right] = \mathbb{E}\left[\prod_{i=0}^{p-1} g_i(K_{U_i})\right] \mu(g_p),$$

de donde se concluye la demostración al usar la hipótesis de inducción. ■

Es de remarcar que en la demostración realizada, el hecho de que  $T < \infty$  con probabilidad uno se sigue de la misma construcción de  $\theta$ .

#### 4. Corolario

Sea  $\{S_n : n \geq 0\}$  una caminata aleatoria sobre  $\mathbb{Z}$  con valor inicial 0 y distribución de saltos dada por  $\nu$ , donde  $\nu(k) = \mu(k+1)$  para cada  $k \geq -1$ . Definamos

$$T = \inf\{n \geq 1 : S_n = -1\}.$$

Entonces el camino de Łukasiewicz asociado a un árbol de Galton–Watson  $\theta$  con distribución de reproducción  $\mu$  tiene la misma distribución que  $(S_0, S_1, \dots, S_T)$ . En particular  $\#\theta$  y  $T$  tienen la misma distribución.

Demostración: Sea  $\{M_n : n \geq 1\}$  una sucesión de variables aleatorias como en la proposición 3 y sea  $T = \{n \geq 1 : M_1 + \dots + M_n < n\}$ . Definamos  $S_0 = 0$  y  $S_n = \sum_{i=1}^n (M_i - 1)$  para  $n \geq 1$ . Entonces  $\{S_n : n \geq 0\}$  es una caminata aleatoria sobre  $\mathbb{Z}$  con la distribución de saltos requerida. En adición, por cómo se definió es igual en distribución al camino de Łukasiewicz asociado a  $\theta$ . Finalmente, que  $T$  y  $\#\theta$  tienen la misma distribución se sigue de

$$T = \inf\{n \geq 1 : S_n = -1\} = \inf\{n \geq M_1 + \dots + M_n < n\}. \quad \blacksquare$$

### 1.3. Convergencia al movimiento browniano

El objetivo de este apartado es demostrar la convergencia débil de las funciones de altura asociadas a un bosque de árboles de Galton–Watson a un movimiento browniano reflejado. A lo largo de la sección se esclarecerá el objetivo, ya que se definirán los objetos necesarios.

Para lograr nuestro cometido supondremos que  $\mu$  es una distribución de reproducción crítica con varianza finita  $\sigma^2 > 0$ . Diremos que una sucesión  $\{\theta_n : n \geq 1\}$  es un bosque de árboles de Galton–Watson si ellos son independientes y para cada  $i$ ,  $\theta_i$  es un GWT( $\mu$ ). Recordando que a cada  $\theta_i$  le podemos asociar su función de altura  $\{h_{\theta_i}(n) : n \in \llbracket 0, \#\theta_i - 1 \rrbracket\}$ , es posible definir el proceso de altura  $\{H_n : n \geq 0\}$  de un bosque concatenando las funciones de altura  $h_{\theta_1}, h_{\theta_2}, \dots$  de los árboles como sigue:

$$H_n = h_{\theta_i} \left( n - \sum_{j=1}^{i-1} \#\theta_j \right) \quad \text{si } n \in \left[ \sum_{j=1}^{i-1} \#\theta_j, \sum_{j=1}^i \#\theta_j - 1 \right].$$

Debido a que  $h_{\theta_i}(n) = 0$  precisamente cuando  $n = 0$ , aunado a que las funciones de altura caracterizan a los árboles, es claro que el proceso de altura da una caracterización de la sucesión de árboles, pues al definir  $T_0 = 0$  y

$$T_i = \inf\{j > T_{i-1} : H_j = 0\}$$

tendremos que

$$h_{\theta_i}(n) = H_{n+T_{i-1}} \quad \text{para } n \in \llbracket 0, T_i - T_{i-1} - 1 \rrbracket.$$

A continuación veremos una proposición que será importante en la demostración del teorema principal de la sección, el teorema 2.

### 1. Proposición

Para cada  $n \geq 0$  se cumple que

$$H_n = \#\left\{k \in \llbracket 0, n-1 \rrbracket : S_k = \inf_{k \leq j \leq n} S_j\right\},$$

donde  $\{S_n : n \geq 0\}$  es una caminata aleatoria con distribución de saltos dada como en el corolario 2.4.

Demostración: Por el corolario 2.4 sabemos que el camino de Łukasiewicz de un árbol  $\theta$  tiene la misma distribución que  $(S_0, \dots, S_T)$  donde  $T = \inf\{n \geq 1 : S_n = -1\}$ . De ello y la proposición 1.4 se obtiene el resultado. ■

Para enunciar el resultado principal recordamos que un movimiento browniano reflejado, que comienza en el origen, es el valor absoluto de un movimiento browniano estándar comenzado en el origen.

### 2. Teorema

Sea  $\{\theta_n : n \geq 1\}$  una colección independiente de árboles de Galton–Watson con distribución de reproducción  $\mu$  y sea  $\{H_n : n \geq 0\}$  el proceso de altura asociado. Entonces

$$\left\{\frac{1}{\sqrt{p}}H_{\lfloor pt \rfloor} : t \geq 0\right\} \Rightarrow \left\{\frac{2}{\sigma}Y_t : t \geq 0\right\} \quad \text{en } \mathbb{D}(\mathbb{R}_+, \mathbb{R}_+) \text{ conforme } p \rightarrow \infty,$$

con  $\gamma$  un movimiento browniano reflejado.

Primero estableceremos la convergencia de las distribuciones finito-dimensionales de los procesos, para lo cual introduciremos procesos auxiliares y estableceremos lemas auxiliares que serán de utilidad en la prueba.

Consideremos  $S = \{S_n : n \geq 0\}$  como en la proposición 1 y notemos que la distribución de saltos  $\nu$  tiene media cero y varianza  $\sigma^2$ , lo cual implica que  $S$  es recurrente. Adicionalmente usaremos la notación

$$M_n = \sup_{0 \leq k \leq n} S_k \quad \text{y} \quad I_n = \inf_{0 \leq k \leq n} S_k.$$

En vista del principio de invarianza de Donsker, ver el teorema 14.1 de Billingsley (1999, pág. 146),

$$\left\{\frac{1}{\sqrt{p}}S_{\lfloor pt \rfloor} : t \geq 0\right\} \Rightarrow \{\sigma B_t : t \geq 0\} \quad \text{conforme } p \rightarrow \infty,$$

donde  $B$  es un movimiento browniano estándar iniciado en el origen. En adición, para cada  $n \in \mathbb{N}_0$ , tomamos la caminata aleatoria en tiempo invertido

$$\widehat{S}^n = \{\widehat{S}_k^n : k \in \llbracket 0, n \rrbracket\} \quad \text{mediante} \quad \widehat{S}_k^n = S_n - S_{(n-k)_+}.$$

Por construcción sabemos que  $\{S_n : k \in \llbracket 0, n \rrbracket\} \stackrel{d}{=} \{\widehat{S}_k^n : k \in \llbracket 0, n \rrbracket\}$ . En otro sentido, observemos que  $\widehat{S}_k^n = \sup_{0 \leq j \leq k} \widehat{S}_j^n$  sucede precisamente cuando

$$S_{n-k} = \inf_{0 \leq j \leq k} S_{n-j} = \inf_{n-k \leq j \leq n} S_j,$$

equivalencia que implica la igualdad

$$\#\left\{k \in \llbracket 0, n-1 \rrbracket : S_k = \inf_{k \leq j \leq n} S_j\right\} = \#\left\{k \in \llbracket 1, n \rrbracket : \hat{S}_k^n = \sup_{0 \leq j \leq k} \hat{S}_j^n\right\}.$$

De esta igualdad, la igualdad en distribución descrita previamente y la proposición 1 obtenemos  $H_n \stackrel{d}{=} K_n$  si definimos

$$K_n = \#\{k \in \llbracket 1, n \rrbracket : S_k = M_k\}.$$

Más aún, usando argumentos similares deducimos, con mayor generalidad, que

$$(M_n, K_n) \stackrel{d}{=} (S_n - I_n, H_n).$$

### 3. Lema

Definamos la sucesión de tiempos de paro  $\{T_n : n \geq 0\}$  de manera inductiva, poniendo  $T_0 = 0$  y para cada  $j \geq 1$ ,

$$T_j = \inf\{n > T_{j-1} : S_n = M_n\}.$$

Entonces las variables aleatorias  $\{S_{T_j} - S_{T_{j-1}} : j \geq 1\}$  conforman una colección independiente e idénticamente distribuidas con

$$\mathbb{P}(S_{T_1} = k) = v([k, \infty)), \quad k \geq 0.$$

Demostración: Que  $\{S_{T_i} - S_{T_{i-1}} : i \geq 1\}$  conforma una colección independiente e idénticamente distribuida se sigue de la propiedad fuerte de Markov de  $S$ , de modo que únicamente hay que calcular la distribución de  $S_{T_1}$  para concluir.

En este sentido, primero observemos que para cualquier  $i \in \mathbb{Z}$ ,

$$\sum_{j \in \mathbb{Z}} v(i - j) = \sum_{j \leq i+1} v(i - j) = \sum_{k \geq -1} v(k) = 1$$

implica que la medida de contar sobre  $\mathbb{Z}$  es una medida invariante de la cadena recurrente  $S$ . Entonces, por el teorema 1.7.5 de Norris (1998), al definir el tiempo de paro  $R_0 = \inf\{n \geq 1 : S_n = 0\}$ , para cada  $i \in \mathbb{Z}$  tendremos que

$$\mathbb{E}\left[\sum_{n=0}^{R_0-1} \mathbb{1}_{\{S_n=i\}}\right] = 1.$$

Puesto que la caminata comienza en 0,  $T_1 \leq R_0$  y  $S_i > 0$  para cada  $i \in \llbracket T_1, R_0 - 1 \rrbracket$ . En consecuencia, para cada  $i \leq 0$ ,

$$\mathbb{E}\left[\sum_{n=0}^{T_1-1} \mathbb{1}_{\{S_n=i\}}\right] = 1,$$

mientras que para  $i > 0$ ,

$$\mathbb{E}\left[\sum_{n=0}^{T_1-1} \mathbb{1}_{\{S_n=i\}}\right] = 0.$$

Por lo tanto, para cualquier función  $g : \mathbb{Z} \rightarrow \mathbb{R}_+$ ,

$$(1) \quad \mathbb{E}\left[\sum_{n=0}^{T_1-1} g(S_n)\right] = \sum_{i \in \mathbb{Z}} \mathbb{E}\left[\sum_{n=0}^{T_1-1} g(S_n) \mathbb{1}_{\{S_n=i\}}\right] = \sum_{i \in \mathbb{Z}} g(i) \mathbb{E}\left[\sum_{n=0}^{T_1-1} \mathbb{1}_{\{S_n=i\}}\right] = \sum_{i \leq 0} g(i).$$

Consideremos  $f : \mathbb{Z} \rightarrow \mathbb{R}_+$  y  $\mathcal{F}_n = \sigma(S_k : k \leq n)$ . Al observar que  $T_1 = k + 1$  si y solo si  $T_1 > k$  y  $S_{k+1} \geq 0$  obtenemos la primera igualdad en

$$\begin{aligned} \mathbb{E}[f(S_{T_1})] &= \sum_{k \geq 0} \mathbb{E} \left[ \mathbb{1}_{\{k < T_1\}} f(S_{k+1}) \mathbb{1}_{\{S_{k+1} \geq 0\}} \right] \\ &= \sum_{k \geq 0} \mathbb{E} \left[ \mathbb{1}_{\{k < T_1\}} \mathbb{E}[f(S_k + (S_{k+1} - S_k)) \mathbb{1}_{\{S_k + (S_{k+1} - S_k) \geq 0\}} \mid \mathcal{F}_k] \right] \\ &= \sum_{k \geq 0} \mathbb{E} \left[ \mathbb{1}_{\{k < T_1\}} \sum_{j \geq -1} v(j) f(S_k + j) \mathbb{1}_{\{S_k + j \geq 0\}} \right] \\ &= \sum_{j \geq -1} v(j) \mathbb{E} \left[ \sum_{k=0}^{T_1-1} f(S_k + j) \mathbb{1}_{\{S_k + j \geq 0\}} \right] \\ &= \sum_{j \geq -1} v(j) \sum_{i \leq 0} f(i + j) \mathbb{1}_{\{i + j \geq 0\}} = \sum_{i \geq 0} \sum_{j \geq i} v(j) f(j - i) \mathbb{1}_{\{j - i \geq 0\}}, \end{aligned}$$

mientras que la tercera igualdad se da por la propiedad de Markov y la quinta por (1). Al usar el cambio de variables  $(m, j) = (j - i, j)$  obtenemos

$$\sum_{i \geq 0} \sum_{j \geq i} v(j) f(j - i) \mathbb{1}_{\{j - i \geq 0\}} = \sum_{m \geq 0} \sum_{j \geq m} v(j) f(m),$$

y por consiguiente

$$\mathbb{E}[f(S_{T_1})] = \sum_{m \geq 0} f(m) \sum_{j \geq m} v(j) = \sum_{m \geq 0} f(m) v([m, \infty)),$$

lo cual nos dice que la distribución de  $S_{T_1}$  es la requerida. ■

Debido al supuesto que realizamos sobre  $\mu$ , que es crítica y tiene varianza  $\sigma^2$ , observamos que

$$\begin{aligned} \mathbb{E}[S_{T_1}] &= \sum_{k \geq 0} k v([k, \infty)) = \sum_{k \geq 0} \sum_{j \geq k} k v(j) = \frac{1}{2} \sum_{j \geq 0} j(j+1) v(j) \\ &= \frac{1}{2} \sum_{j \geq 1} j(j-1) \mu(j) = \frac{1}{2} \sum_{j \geq 0} (j-1)^2 \mu(j) + \frac{1}{2} \sum_{j \geq 0} (j-1) \mu(j) = \frac{1}{2} \sigma^2. \end{aligned}$$

Usaremos este resultado en la prueba del siguiente lema, que será vital para la demostración de la convergencia de las distribuciones finito-dimensionales del teorema 2.

#### 4. Lema

Se tiene que

$$\frac{H_n}{S_n - I_n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \frac{2}{\sigma^2},$$

donde  $\xrightarrow{\mathbb{P}}$  representa convergencia en probabilidad.

Demostración: De las definiciones de  $M_n$ ,  $T_k$  y  $K_n$  deducimos que

$$(2) \quad M_n = \sum_{T_j \leq n} (S_{T_k} - S_{T_{k-1}}) = \sum_{k=1}^{K_n} (S_{T_k} - S_{T_{k-1}}).$$

Ahora, es fácil verificar que  $K_n \rightarrow \infty$  cuando  $n \rightarrow \infty$  sucede con probabilidad uno. En efecto, como  $K_n \leq K_{n+1}$ , basta probar la convergencia en probabilidad, para lo cual basta notar que

$$\mathbb{P}(K_n \geq j) = \mathbb{P}(T_j \leq n),$$

para  $j \geq 1$  fija, y en consecuencia

$$\lim_{n \rightarrow \infty} \mathbb{P}(K_n \geq j) = \mathbb{P}(T_j < \infty) = 1,$$

donde la última igualdad se da por tratarse de una cadena recurrente. Entonces, por la ley fuerte de grandes números obtenemos que con probabilidad uno,

$$\frac{M_n}{K_n} \xrightarrow[n \rightarrow \infty]{} \mathbb{E}[S_{T_1}] = \frac{\sigma^2}{2}.$$

Así, al usar la identidad distribucional  $(M_n, K_n) \stackrel{d}{=} (S_n - I_n, H_n)$  concluimos que

$$\frac{S_n - I_n}{H_n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \frac{\sigma^2}{2},$$

con lo cual obtenemos el lema. ■

Procedemos a la primera parte de la demostración del teorema 2.

Demostración de la convergencia finito-dimensional en el teorema 2: Del principio de invarianza de Donker tenemos

$$\frac{1}{\sqrt{p}} \left( S_{\lfloor pt_1 \rfloor} - I_{\lfloor pt_1 \rfloor}, \dots, S_{\lfloor pt_m \rfloor} - I_{\lfloor pt_m \rfloor} \right) \Rightarrow \sigma \left( B_{t_1} - \inf_{0 \leq s \leq t_1} B_s, \dots, B_{t_m} - \inf_{0 \leq s \leq t_m} B_s \right)$$

conforme  $p \rightarrow \infty$  para cualesquiera  $0 \leq t_1 \leq \dots \leq t_m$ . Así, por el lema 4,

$$\frac{1}{\sqrt{p}} \left( H_{\lfloor pt_1 \rfloor}, \dots, H_{\lfloor pt_m \rfloor} \right) \Rightarrow \frac{2}{\sigma} \left( B_{t_1} - \inf_{0 \leq s \leq t_1} B_s, \dots, B_{t_m} - \inf_{0 \leq s \leq t_m} B_s \right).$$

Por la identidad de Lévy, teorema 2.3 o teorema 9.14 de Le Gall (2016, pág. 252), sabemos que

$$\gamma_t = B_t - \inf_{0 \leq s \leq t} B_s$$

define un movimiento browniano reflejado. ■

Ahora procederemos a demostrar la convergencia funcional del teorema 2. Siguiendo las notas de Le Gall (2005), únicamente consideraremos el caso donde  $\mu$  tiene momentos exponenciales “pequeños”, es decir que supondremos la existencia de  $\lambda > 0$  tal que

$$\sum_{k \geq 0} e^{\lambda k} \mu(k) < \infty.$$

Para establecer el resultado bajo este supuesto, enunciaremos y demostraremos dos lemas que se usarán en la demostración.

**5. Lema**

Sea  $\{Y_n : n \geq 1\}$  una sucesión de variables aleatorias independientes e idénticamente distribuidas con  $\mathbb{E}[Y_1] = 0$ . Supongamos que existe  $\lambda > 0$  para el cual  $\mathbb{E}[\exp\{\lambda|Y_1|\}] < \infty$ . Entonces, para cada  $\alpha > 0$ , existe  $N$  para la cual se cumple

$$\mathbb{P}(|Y_1 + \cdots + Y_l| > n^{\alpha+1/2}) \leq \exp\{-n^{\alpha/2}\}$$

siempre que  $n \geq N$  y  $l \in \llbracket 1, n \rrbracket$ .

Demostración: La hipótesis implica la existencia de  $\mathbb{E}[e^{\lambda Y_1}]$  para  $\lambda$  suficientemente pequeña, lo cual implica, de acuerdo a Billingsley (1995, pág. 278), que  $Y_1$  tiene momentos finitos de todos los órdenes y

$$\mathbb{E}[e^{\lambda Y_1}] = \sum_{k \geq 0} \frac{\lambda^k}{k!} \mathbb{E}[Y_1^k] = 1 + c\lambda^2 + o(\lambda^2)$$

conforme  $\lambda \rightarrow 0$ , donde hemos puesto  $c = \text{Var}(Y_1)/2$ . En consecuencia, existe una constante  $C > c$  para la cual se cumple

$$\mathbb{E}[e^{\lambda Y_1}] \leq 1 + C\lambda^2 \leq e^{C\lambda^2}$$

para  $\lambda > 0$  suficientemente pequeña.

Por consiguiente, para  $\lambda > 0$  suficientemente pequeña tendremos, gracias a la desigualdad de Markov,

$$\mathbb{P}(Y_1 + \cdots + Y_l > n^{\alpha+1/2}) \leq e^{-\lambda n^{\alpha+1/2}} \mathbb{E}[e^{\lambda(Y_1 + \cdots + Y_l)}] \leq \exp\{-\lambda n^{\alpha+1/2} + Cn\lambda^2\}.$$

Reemplazando  $Y_i$  por  $-Y_i$  y poniendo  $\lambda = n^{-1/2}$  para  $n$  suficientemente grande obtenemos

$$\begin{aligned} \mathbb{P}(|Y_1 + \cdots + Y_l| > n^{\alpha+1/2}) &\leq 2 \exp\{-n^{\alpha} + C\} \\ &= \exp\{-n^{\alpha/2}\} (2 \exp\{C + n^{\alpha/2} - n^{\alpha}\}). \end{aligned}$$

Concluimos observando que para  $n \geq 2^{2/\alpha} \vee (C + \log 2)^{2/\alpha}$ ,

$$2 \exp\{C + n^{\alpha/2}\} \leq \exp\{n^{\alpha}\}. \quad \blacksquare$$

**6. Lema**

Sea  $\varepsilon \in (0, 1/4)$ . Existen  $\varepsilon' > 0$  y un entero  $N \geq 1$  tales que para cualquier elección de  $n \geq N$  y  $l \in \llbracket 0, n \rrbracket$ ,

$$\mathbb{P}\left(\left|M_l - \frac{\sigma^2}{2} K_l\right| > n^{\varepsilon+1/4}\right) < \exp\{-n^{\varepsilon'}\}.$$

Demostración: Tomemos  $\alpha \in (0, \varepsilon/2)$  y pongamos  $m_n = \lfloor n^{\alpha+1/2} \rfloor$ . Entonces para cada  $l \in \llbracket 0, n \rrbracket$  tendremos

$$\mathbb{P}\left(\left|M_l - \frac{\sigma^2}{2} K_l\right| > n^{\varepsilon+1/4}\right) \leq \mathbb{P}(K_l > m_n) + \mathbb{P}\left(\left|M_l - \frac{\sigma^2}{2} K_l\right| > n^{\varepsilon+1/4}, K_l \leq m_n\right).$$

De (2) podemos reescribir

$$M_l - \frac{\sigma^2}{2} K_l = \sum_{j=1}^{K_l} \left( (S_{T_j} - S_{T_{j-1}}) - \frac{\sigma^2}{2} \right),$$

con lo cual obtenemos las contenciones

$$\begin{aligned} & \left\{ \left| M_l - \frac{\sigma^2}{2} K_l \right| > n^{\varepsilon+1/4}, K_l \leq m_n \right\} \\ & \subset \left\{ \left| \sum_{j=1}^{K_l} \left( (S_{T_j} - S_{T_{j-1}}) - \frac{\sigma^2}{2} \right) \right| > n^{\varepsilon+1/4}, K_l \leq m_n \right\} \\ & \subset \left\{ \sup_{0 \leq k \leq m_n} \left| \sum_{j=1}^k \left( (S_{T_j} - S_{T_{j-1}}) - \frac{\sigma^2}{2} \right) \right| > n^{\varepsilon+1/4}, K_l \leq m_n \right\} \\ & \subset \left\{ \sup_{0 \leq k \leq m_n} \left| \sum_{j=1}^k \left( (S_{T_j} - S_{T_{j-1}}) - \frac{\sigma^2}{2} \right) \right| > n^{\varepsilon+1/4} \right\}. \end{aligned}$$

Por ende, usando la monotonía de probabilidad y el hecho de que  $m_n \leq n^{\alpha+1/2}$ , que en particular implica

$$m_n^{\varepsilon+1/2} \leq n^{\varepsilon\alpha+(\varepsilon+\alpha)/2+1/4} \leq n^{\varepsilon(\varepsilon/2+\varepsilon/4+1/2)+1/4} \leq n^{\varepsilon+1/4},$$

obtenemos la primera desigualdad en

$$\begin{aligned} & \mathbb{P} \left( \left| M_l - \frac{\sigma^2}{2} K_l \right| > n^{\varepsilon+1/4}, K_l \leq m_n \right) \\ & \leq \mathbb{P} \left( \sup_{0 \leq k \leq m_n} \left| \sum_{j=1}^k \left( (S_{T_j} - S_{T_{j-1}}) - \frac{\sigma^2}{2} \right) \right| > m_n^{\varepsilon+1/2} \right) \leq m_n \exp \{ -m_n^{\varepsilon/2} \}, \end{aligned}$$

donde la última cota se vale para  $n$  suficientemente grande en vista del lema 5. Para poder usar el lema 5 es necesario que  $S_{T_1}$  tenga momentos exponenciales pequeños, lo cual sucede en vista de las igualdades

$$\sum_{k \geq 0} e^{\lambda k} \nu([k, \infty)) = \sum_{j \geq 0} \mu(j+1) \sum_{k=0}^j e^{\lambda k} = \sum_{j \geq 0} \mu(j+1) \frac{e^{\lambda(j+1)} - 1}{e^{\lambda} - 1}$$

y que supusimos que  $\mu$  tiene momentos exponenciales pequeños.

Ahora acotaremos el término  $\mathbb{P}(K_l > m_n)$  notando primero que por monotonía  $\mathbb{P}(K_l > m_n) \leq \mathbb{P}(K_n > m_n)$ . Ahora, por definición de  $K_n$  y  $T_j$  se tiene que  $K_n > m_n$  implica  $S_{T_{m_n}} \leq M_n$ ; por tanto

$$\mathbb{P}(K_l > m_n) \leq \mathbb{P}(S_{T_{m_n}} \leq M_n) \leq \mathbb{P}(S_{T_{m_n}} \leq n^{(\alpha+1)/2}) + \mathbb{P}(M_n > n^{(\alpha+1)/2}).$$

Por definición de  $M_n$  tenemos la igualdad  $\{M_n > n^{(\alpha+1)/2}\} = \bigcup_{j=1}^n \{S_j > n^{(\alpha+1)/2}\}$  que nos permitirá obtener las cotas en

$$\begin{aligned} \mathbb{P}(M_n > n^{(\alpha+1)/2}) & \leq \sum_{j=1}^n \mathbb{P}(S_j > n^{(\alpha+1)/2}) \leq n \sup_{1 \leq j \leq n} \mathbb{P}(S_j > n^{(\alpha+1)/2}) \\ & \leq n \exp \{ -n^{\alpha/4} \}, \end{aligned}$$

donde la última cota es válida para  $n$  suficientemente grande en virtud del lema 5. Por su parte, notando que para  $n$  suficientemente grande  $n^{(\alpha+1)/2} \leq \sigma^2 m_n / 4$  obtenemos

$$\begin{aligned} \mathbb{P}(S_{T_{m_n}} \leq n^{(\alpha+1)/2}) &= \mathbb{P}\left(S_{T_{m_n}} - \frac{\sigma^2}{2} m_n \leq n^{(\alpha+1)/2} - \frac{\sigma^2}{2} m_n\right) \\ &\leq \mathbb{P}\left(\left|\sum_{j=1}^{m_n} (S_{T_j} - S_{T_{j-1}} - \frac{\sigma^2}{2})\right| \geq \frac{\sigma^2}{2} m_n - n^{(\alpha+1)/2}\right) \\ &\leq \mathbb{P}\left(\left|\sum_{j=1}^{m_n} (S_{T_j} - S_{T_{j-1}} - \frac{\sigma^2}{2})\right| > \frac{\sigma^2}{8} n^{\alpha+1/2}\right) \leq \exp\{-n^{\alpha/2}\}. \end{aligned}$$

Para la penúltima desigualdad se usaron las desigualdades, válidas para  $n$  suficientemente grande,

$$\frac{\sigma^2}{2} m_n - n^{(\alpha+1)/2} \geq \frac{\sigma^2}{4} m_n \geq \frac{\sigma^2}{8} n^{\alpha+1/2}.$$

Así, puesto que para  $n$  suficientemente grande  $m_n \geq n^{1/2}$ , de las cotas deducidas se sigue que

$$\begin{aligned} \mathbb{P}\left(\left|M_l - \frac{\sigma^2}{2} K_l\right| > n^{\varepsilon+1/4}\right) &\leq m_n \exp\{-m_n^{\varepsilon/2}\} + n \exp\{-n^{\alpha/4}\} + \exp\{-n^{\alpha/2}\} \\ &\leq 3n \exp\{-n^{\alpha/4}\}. \end{aligned}$$

De esta última desigualdad se obtiene el resultado. ■

Concluimos la sección con la demostración de la convergencia funcional en el teorema 2 bajo el supuesto de la existencia de momentos pequeños.

Demostración de la convergencia funcional del teorema 2: Aplicaremos el lema 6 con  $\varepsilon = 1/8$  para obtener la existencia de  $\varepsilon' > 0$  tal que

$$\mathbb{P}\left(\left|S_l - I_l - \frac{\sigma^2}{2} H_l\right| > n^{3/8}\right) = \mathbb{P}\left(\left|M_l - \frac{\sigma^2}{2} K_l\right| > n^{3/8}\right) < \exp\{-n^{\varepsilon'}\}$$

para  $n$  suficientemente grande y  $l \in \llbracket 0, n \rrbracket$ . De lo anterior,

$$\mathbb{P}\left(\sup_{0 \leq l \leq n} \left|S_l - I_l - \frac{\sigma^2}{2} H_l\right| > n^{3/8}\right) < n \exp\{-n^{\varepsilon'}\}.$$

Dado un entero  $A \geq 1$  fijo, la desigualdad previa implica, para  $p$  suficientemente grande,

$$\begin{aligned} (3) \quad \mathbb{P}\left(\sup_{0 \leq t \leq A} \left|\frac{S_{\lfloor pt \rfloor} - I_{\lfloor pt \rfloor}}{\sqrt{p}} - \frac{\sigma^2}{2} \frac{H_{\lfloor pt \rfloor}}{\sqrt{p}}\right| > \frac{A^{3/8}}{p^{1/8}}\right) \\ \leq \mathbb{P}\left(\sup_{0 \leq t \leq A} \left|S_{\lfloor pt \rfloor} - I_{\lfloor pt \rfloor} - \frac{\sigma^2}{2} H_{\lfloor pt \rfloor}\right| > (Ap)^{3/8}\right) < (Ap) \exp\{-(Ap)^{\varepsilon'}\}. \end{aligned}$$

En consecuencia

$$\sum_{p \geq 1} \mathbb{P}\left(\sup_{0 \leq t \leq A} \left|\frac{S_{\lfloor pt \rfloor} - I_{\lfloor pt \rfloor}}{\sqrt{p}} - \frac{\sigma^2}{2} \frac{H_{\lfloor pt \rfloor}}{\sqrt{p}}\right| > \frac{A^{3/8}}{p^{1/8}}\right) < \infty,$$

lo cual implica, a través del lema de Borel–Cantelli, que

$$\mathbb{P}\left(\sup_{0 \leq t \leq A} \left|\frac{S_{\lfloor pt \rfloor} - I_{\lfloor pt \rfloor}}{\sqrt{p}} - \frac{\sigma^2}{2} \frac{H_{\lfloor pt \rfloor}}{\sqrt{p}}\right| > \frac{A^{3/8}}{p^{1/8}} \text{ i.o.}\right) = 0.$$

Es claro que lo anterior implica que, con probabilidad uno,

$$\sup_{0 \leq t \leq A} \left| \frac{S_{[pt]} - I_{[pt]}}{\sqrt{pt}} - \frac{\sigma^2}{2} \frac{H_{[pt]}}{\sqrt{pt}} \right| \xrightarrow{p \rightarrow \infty} 0.$$

Luego, por el lema 3 de la sección 16 de Billingsley (1999, pág. 173), el teorema 3.1 de la misma referencia (pág. 27) y lo demostrado, el resultado se obtiene si

$$\left\{ \frac{S_{[pt]} - I_{[pt]}}{\sqrt{p}} : t \geq 0 \right\} \Rightarrow \left\{ \sigma \left( B_t - \inf_{0 \leq s \leq t} B_s \right) : t \geq 0 \right\}$$

conforme  $p \rightarrow \infty$ . Puesto que tal convergencia es una consecuencia del principio de invarianza de Donsker, hemos concluido. ■

# Tiempos locales y excursiones

# 2

En este capítulo introducimos la noción de *tiempo local* para una colección de procesos de Markov conocidos como *procesos de Hunt*. Esto nos permitirá describir la evolución de un proceso de Markov en términos de excursiones, que son las trayectorias del proceso entre visitas a un punto del espacio. Las ideas que se desarrollan siguen de cerca la exposición de los capítulos 2 y 3 de las notas «Local Times and Excursion Theory» de Pardo Millán (s.f.-a).

## 2.1. Procesos de Markov

En esta sección introducimos la noción de proceso de Hunt, el cual es la subclase de procesos de Markov con la cual estaremos trabajando en apartados posteriores. La referencia principal es Pardo Millán (s.f.-a) y algunas referencias suplementarias son la sección iv.9 de Blumenthal y Gettoor (1968) y la sección 47 de Sharpe (1988). En particular, las definiciones de proceso estándar y de proceso de Hunt se adaptan del libro de Blumenthal y Gettoor (1968).

A lo largo de esta sección consideraremos que  $(\Omega, \mathcal{F}, \mathbb{P})$  es un espacio de probabilidad, que  $E$  es un espacio métrico localmente compacto con un punto distinguido  $0$ . Además consideramos un proceso estocástico  $X = \{X_t : t \geq 0\}$  con valores en  $E$ , trayectorias continuas por la derecha (càd) y tal que  $\mathbb{P}(X_0 = 0) = 0$ . Supondremos que este proceso está adaptado a una filtración  $\{\mathcal{F}_t : t \geq 0\}$  que satisface las hipótesis usuales<sup>1</sup>. Finalmente, supondremos que existe una familia de medidas de probabilidad  $\{\mathbb{P}_x : x \in E\}$  que satisface:

**(PFM)** Para cualquier tiempo de paro finito  $\tau$ , el proceso desplazado

$$X \circ \theta_\tau := \{X_{\tau+t} : t \geq 0\},$$

bajo la ley condicional  $\mathbb{P}(\cdot \mid X_\tau = x)$ , es independiente de  $\mathcal{F}_\tau$  y tiene ley  $\mathbb{P}_x$ .

Aquellos procesos estocásticos que cumplen la propiedad (PFM) son conocidos como *procesos de Markov fuerte*.

En ciertas ocasiones será de utilidad considerar agregar un punto  $\Delta$  a  $E$ , el cual será un punto “al infinito” si  $E$  no es compacto, y un punto aislado en caso que  $E$  sea compacto. Una vez hecho esto, definiremos el *tiempo de vida* de  $X$ ,

$$\zeta = \inf\{t > 0 : X_t = \Delta\}.$$

En lo subsecuente requeriremos que  $E$  tenga base contable, lo cual es de hecho equivalente a que  $E$  sea polaco, ver el teorema 5.3 de Kechris (1995, pág. 59), es decir que  $E$  es un espacio métrico completo y separable.

<sup>1</sup>Ser completa y continua por la derecha.

**1. Definición** (Proceso estándar)

Diremos que un proceso de Markov fuerte  $X$  es un *proceso estándar* si él es *cuasicontinuo por la izquierda*: Dada cualquier sucesión creciente  $\{T_n : n \geq 1\}$  de tiempos de paro con respecto a  $\{\mathcal{F}_t : t \geq 0\}$  que converja a  $T$ , se tiene  $X_{T_n} \rightarrow X_T$  con probabilidad uno sobre el evento  $\{T < \zeta\}$ .

La condición de cuasicontinuidad por la izquierda se puede hacer más restrictiva; en efecto, diremos que un proceso  $X$  es cuasicontinuo por la izquierda en  $[0, \infty)$  si  $X_{T_n} \rightarrow X_T$  se da con probabilidad uno en el evento  $\{T < \infty\}$ . Este es el tipo de cuasicontinuidad con la cual trabajaremos de ahora en adelante.

**2. Definición** (Proceso de Hunt)

Un proceso estándar  $X$  es un *proceso de Hunt* si es cuasicontinuo por la izquierda en  $[0, \infty)$ .

**3. Proposición**

Sean  $B$  un subconjunto abierto o cerrado de  $E$  y  $X$  un proceso de Hunt. Entonces

$$T_B = \inf\{t \geq 0 : X_t \in B\}$$

define un tiempo de paro con respecto a  $\{\mathcal{F}_t : t \geq 0\}$ .

Demostración: Supongamos primero que  $B$  es abierto. Por la continuidad por la derecha de  $X$ , para cada  $t > 0$  tendremos que

$$\{T_B < t\} = \bigcup_{s \in [0, t) \cap \mathbb{Q}} \{X_s \in B\} \in \mathcal{F}_t,$$

y en consecuencia  $\{T_B \leq t\} \in \mathcal{F}_{t+} = \mathcal{F}_t$  por tratarse de una filtración continua por la derecha. Así, hemos demostrado que  $T_B$  es un tiempo de paro.

En caso de que  $B$  sea cerrado, notamos que la  $\varepsilon$ -vencindad de  $B$ , definida por  $B^\varepsilon = \{x \in E : d_E(x, B) < \varepsilon\}$ , es un conjunto abierto, de modo que  $T_{B^{1/n}}$  es un tiempo de paro. En adición  $T_{B^{1/n}} \leq T_{B^{1/(n+1)}}$ , motivo por el cual  $\{T_{B^{1/n}}\}$  crece a un tiempo de paro  $T'$ . En consecuencia, concluiremos si probamos que  $T' = T_B$  casi seguramente. Para ello primeramente observamos que  $T' \leq T_B$  con probabilidad uno.

Para establecer la desigualdad contraria hacemos notar que debido a la cuasicontinuidad por la izquierda de  $X$ ,  $X_{T_{B^{1/n}}} \rightarrow X_{T'}$  con probabilidad uno en el evento  $\{T' < \infty\}$ . Debido a la continuidad por la derecha de  $X$ ,  $X_{T_{B^{1/(n+1)}}} \in \text{cl}(B^{1/(n+1)}) \subset B^{1/n}$  en el conjunto  $\{T_{B^{1/(n+1)}} < \infty\}$ . Por consiguiente, en el conjunto  $\{T' < \infty\}$ ,  $X_{T'} \in \bigcap_{n \geq 1} B^{1/n} = B$ , donde la igualdad se da por ser  $B$  cerrado. Esto implica  $T_B \leq T'$  en el conjunto  $\{T' < \infty\}$ , lo cual nos da el resultado deseado. ■

**4. Teorema**

Con probabilidad uno, las trayectorias de un proceso de Hunt  $X$  tienen límite por la izquierda. Dicho de otra forma, un proceso de Hunt es càdlàg con probabilidad uno.

Demostración: Dado  $\varepsilon > 0$  definamos el tiempo aleatorio

$$T^{(\varepsilon)} = \inf\{t > 0 : d_E(X_t, 0) > \varepsilon\},$$

el cual es un tiempo de paro en vista de la proposición 3. Definamos los tiempos de paro con respecto a  $\{\mathcal{F}_t : t \geq 0\}$  mediante  $T_0^{(\varepsilon)} = 0$  y para  $n \geq 1$ ,

$$T_n^{(\varepsilon)} = T_{n-1}^{(\varepsilon)} + T^{(\varepsilon)} \circ \theta_{T_{n-1}^{(\varepsilon)}} = \inf\{t > T_{n-1}^{(\varepsilon)} : d_E(X_t, X_{T_{n-1}^{(\varepsilon)}}) > \varepsilon\}.$$

Observemos que en particular  $T_1^{(\varepsilon)} = T^{(\varepsilon)}$ .

Es claro que los tiempos de paro definidos son crecientes en  $n$ , de modo que  $\tau^{(\varepsilon)} = \lim_n T_n^{(\varepsilon)}$  también será un tiempo de paro. Al observar que en el conjunto  $\{T_n < \infty\}$  se cumple  $d_E(X_{T_{n-1}^{(\varepsilon)}}, X_{T_n^{(\varepsilon)}}) \geq \varepsilon$  debido a la continuidad por derecha de las trayectorias de  $X$ , deducimos  $T = \infty$  con probabilidad uno. En caso contrario se violaría la cuasicontinuidad por la izquierda de  $X$ .

Por otra parte, definamos al proceso  $X^{(\varepsilon)}$  a través de

$$X_t^{(\varepsilon)} = \sum_{n \geq 1} X_{T_{n-1}^{(\varepsilon)}} \mathbb{1}_{[T_{n-1}^{(\varepsilon)}, T_n^{(\varepsilon)})}(t).$$

Es claro que  $X^{(\varepsilon)}$  es càdlàg como función de  $t$ . Notando que

$$\sup_{t \in [T_{n-1}^{(\varepsilon)}, T_n^{(\varepsilon)})} d_E(X_t, X_{T_{n-1}^{(\varepsilon)}}) \leq \varepsilon$$

obtenemos

$$\sup_{t \in \mathbb{R}_+} d_E(X_t, X_t^{(\varepsilon)}) \leq \varepsilon \text{ casi seguramente.}$$

Por lo tanto, con probabilidad uno,  $X$  es el límite uniforme de funciones càdlàg, con lo cual se sigue el resultado. ■

Dado un tiempo de paro  $T$  con respecto a  $\{\mathcal{F}_t : t \geq 0\}$ , se tiene que  $\{T = 0\} \in \mathcal{F}_0$  y por lo tanto, debido a la ley cero-uno de Blumenthal, ver la proposición 5.7 de Blumenthal y Gettoor (1968), para cada  $x \in E$ ,  $\mathbb{P}_x(T = 0)$  es cero o uno. En caso de que  $\mathbb{P}_x(T = 0) = 1$  diremos que  $x$  es *regular* para  $T$  y diremos que es *irregular* para  $T$  si la probabilidad es cero. Esto nos lleva a la siguiente definición.

### 5. Definición

Para un conjunto  $B$  abierto o cerrado,  $x$  es *regular* (*irregular*) para  $B$  si es regular (irregular resp.) para  $T_B$ .

En adición, diremos que  $x \in E$  es *regular* si es regular para  $\{x\}$ .

De ahora en adelante supondremos que  $X$  es un proceso de Hunt. Además usaremos que para cualquier tiempo de paro  $T$ , el primer tiempo de retorno al estado inicial tras  $T$ ,

$$R_T = \inf\{t > T : X_t = 0\},$$

es un tiempo de paro. Por la discusión previa a la definición 5,  $\mathbb{P}(R_0 = 0) \in \{0, 1\}$ . Cuando 0 sea regular, introducimos el primer tiempo de salida de 0,

$$F = \inf\{t \geq 0 : X_t \neq 0\},$$

el cual es un tiempo de paro y por ende  $\mathbb{P}(F = 0) \in \{0, 1\}$ . Si  $\mathbb{P}(F = 0) = 0$ , 0 será un *punto de retención* y 0 será un *punto instantáneo* de suceder  $\mathbb{P}(F = 0) = 1$ .

## 2.2. Tiempo local

En esta sección introduciremos la noción de tiempos locales para procesos de Markov, o mejor dicho procesos de Hunt. De manera informal, el tiempo local del punto  $x \in E$  para un proceso  $X$  es un funcional aditivo que crece en el conjunto de tiempos en el cual  $X$  visita  $x$ . Nuevamente nos guiamos principalmente de Pardo Millán (s.f.-a) y en algunas ocasiones seguimos de la sección IV.2 del libro *Lévy Processes* de Bertoin (1996).

A lo largo de la sección supondremos que 0 es regular (para sí mismo) e instantáneo. Estas suposiciones las realizamos en vista de que este caso es el más interesante, pues cuando 0 es irregular o punto de retención, es claro cómo definir el tiempo local (del 0).

Antes de construir el tiempo local, habrá que estudiar algunas propiedades del conjunto de ceros de  $X$ , que denotaremos por  $\mathcal{Z} = \{t \geq 0 : X_t = 0\}$ . Primero notemos que de la continuidad por la derecha de  $X$  se obtiene que si existe una sucesión  $\{t_n : n \geq 1\} \subset \mathcal{Z}$  que decrezca a  $t \geq 0$  entonces  $t \in \mathcal{Z}$ . De esta observación se desprende que todo punto en  $\text{cl}(\mathcal{Z}) \setminus \mathcal{Z}$  es aislado por la derecha. Si a ello le aunamos el hecho de que  $\text{cl}(\mathcal{Z})^c$  es abierto y por ende unión contable de intervalos abiertos disjuntos, y la igualdad  $\mathcal{Z}^c = (\text{cl}(\mathcal{Z}) \setminus \mathcal{Z}) \cup \text{cl}(\mathcal{Z})^c$ , concluimos que  $\mathcal{Z}^c$  es unión disjunta de intervalos de la forma  $(u, v)$  o  $[u, v)$ . Para enunciar el siguiente resultado, recordemos que un conjunto  $\mathcal{A}$  dicese *denso en ninguna parte* si su cerradura tiene interior vacío, es decir  $\text{int}(\text{cl}(\mathcal{A})) = \emptyset$ , y dicese *denso en sí mismo* si no tiene puntos aislados.

### 1. Proposición

*El conjunto  $\mathcal{Z}$  es denso en ninguna parte y denso en sí mismo.*

**Demostración:** Debido a que 0 es instantáneo,  $F = 0$  casi seguramente. Por ende, en el conjunto  $\{R_r < \infty\}$  tenemos  $F \circ \theta_{R_r} = 0$  casi seguramente, y en consecuencia  $R_r \in \text{cl}(\mathcal{Z}^c)$  con probabilidad uno en el mismo evento. Como  $\{R_r : r \in \mathbb{Q}_+\}$  es denso en  $\text{cl}(\mathcal{Z})$  se sigue que  $\text{cl}(\mathcal{Z}) \subset \text{cl}(\mathcal{Z}^c)$ , de modo que  $\text{cl}(\mathcal{Z}^c) = \mathbb{R}_+$  casi seguramente. De las observaciones realizadas previamente, sabemos que  $\mathcal{Z}^c$  es una unión de intervalos de la forma  $(u, v)$  y  $[u, v)$  y por consiguiente  $\text{cl}(\text{int}(\mathcal{Z}^c)) = \text{cl}(\mathcal{Z}^c)$ . Al tomar complementos, obtenemos que con probabilidad uno  $\text{int}(\text{cl}(\mathcal{Z})) = \emptyset$ , es decir que  $\mathcal{Z}$  es denso en ninguna parte.

Veamos ahora que  $\mathcal{Z}$  es denso en sí mismo. Para ello comencemos notando que  $R_0 = 0$  casi seguramente pues 0 es regular. Consecuentemente, en el conjunto  $\{R_r < \infty\}$ ,  $R_0 \circ \theta_{R_r} = 0$  con probabilidad uno. Por el punto anterior, todo punto en  $\mathcal{Z}$  es de la forma  $R_r$  para algún  $r \in \mathbb{Q}_+$ . Esto implica que  $\mathcal{Z}$  es denso en sí mismo. ■

Diremos que un intervalo abierto  $(g, d)$  es un *intervalo de excursión* si es maximal entre los intervalos abiertos en los cuales  $X = 0$ . De manera equivalente,  $(g, d)$  es un intervalo de excursión si  $X_t = 0$  para toda  $t \in (g, d)$ ,  $g \in \text{cl}(\mathcal{Z})$  y  $d \in \text{cl}(\mathcal{Z}) \cup \{\infty\}$ . En este caso  $g$  (*gauche*) es el extremo izquierdo,  $d$  (*droite*) es el extremo derecho y  $\ell = d - g$  es la longitud del intervalo de excursión. A los intervalos de excursión los podemos dotar con un orden total natural, dado por  $(g, d) < (g', d')$  si y sólo si  $g < d \leq g' < d'$ . Una observación interesante es que estos intervalos son aquellos

que aparecen en la representación como unión de intervalos abiertos disjuntos de  $\text{cl}(\mathcal{Z})^c$ .

Ahora consideramos un real  $c > 0$  para el cual se tenga

$$(1) \quad \mathbb{P}(\text{existe al menos una excursión con } \ell > c) > 0.$$

La continuidad por la derecha de  $X$  nos asegura la existencia de alguna constante salvo que el proceso  $X$  sea idénticamente cero, lo cual no sucede en vista de que estamos suponiendo que 0 es instantáneo. La elección de  $c$  es arbitraria y no afectará de forma importante los resultados que se presentarán. Una vez fijada  $c$ , se puede verificar que la probabilidad en (1) es de hecho igual a 1. En efecto, definamos

$$\Lambda_t = \{\text{todas las excursiones con extremo derecho } d < t \text{ tienen longitud } \ell \leq c\}$$

para  $t > 0$  y escojamos  $t$  suficientemente grande de tal manera que  $\mathbb{P}(\Lambda_t) < 1$ . Analizando el tiempo de paro  $R_t$  vemos que es posible que sea infinito, en cuyo caso hay un intervalo de excursión infinito, y en consecuencia tiene un intervalo con longitud mayor que  $c$ . En caso de que sea finito, vemos que en el evento  $\Lambda_{3t}$ ,  $R_t \in (t, 2t)$ , con lo cual obtenemos, aplicando la propiedad fuerte de Markov, la desigualdad  $\mathbb{P}(\Lambda_{3t}) \leq \mathbb{P}(\Lambda_t)^2$ . De manera iterativa obtenemos

$$\mathbb{P}(\Lambda_{3^n t}) \leq \mathbb{P}(\Lambda_t)^{2^n} \text{ para } n \geq 1,$$

de donde deducimos  $\lim_{s \rightarrow \infty} \mathbb{P}(\Lambda_s) = 0$ , dándonos la afirmación.

Para cada elección de  $a > 0$  y  $n \in \mathbb{N}$ , denotamos por  $\ell_n(a)$ ,  $g_n(a)$  y  $d_n(a)$  a la longitud, extremo izquierdo y extremo derecho de la  $n$ -ésima excursión con longitud  $\ell > a$ . En caso de que el número de intervalos de longitud  $\ell > a$  sea menor que  $n$  ponemos  $\ell_n(a) = 0$  y  $g_n(a) = d_n(a) = \infty$ . Notando que

$$d_1(a) = \inf\{t > 0 : X_t = 0 \text{ y } X_{t-s} \neq 0 \text{ para } s \in (0, a]\}$$

vemos que  $d_1(a)$  es un tiempo de paro e inductivamente se deduce que  $d_n(a)$  es un tiempo de paro para cada  $n \geq 1$ .

## 2. Lema

Para todo  $a \in (0, c)$  se cumple  $\mathbb{P}(\ell_1(a) > c) > 0$ .

Demostración: Procedemos por contradicción suponiendo que existe  $a \in (0, c)$  para la cual la probabilidad de interés sea 0. Entonces  $d_1(a) < g_1(c)$  casi seguramente. Aplicando la propiedad de Markov en los tiempos de paro  $d_n(a)$  obtenemos, recursivamente, que  $d_n(a) < g_1(c)$  casi seguramente para  $n \geq 1$ . De lo anterior se sigue  $g_1(c) = \infty$  con probabilidad uno, lo cual es la contradicción requerida. ■

Con lo probado, definimos la función  $\bar{\Pi} : (0, \infty] \rightarrow (0, \infty)$  mediante

$$\bar{\Pi}(a) = \begin{cases} \frac{1}{\mathbb{P}(\ell_1(a) > c)} & \text{si } a \leq c, \\ \mathbb{P}(\ell_1(c) > a) & \text{si } a > c. \end{cases}$$

Por definición de  $\ell_1(c)$  es claro que  $\bar{\Pi}(c) = 1$ . Por otra parte, si  $a < c$  se tendrá

$$\{\ell_1(a) > c\} = \bigcap_{b \in (a, c)} \{\ell_1(b) > c\},$$

mientras que en el caso  $c \leq a$  tendremos

$$\{\ell_1(c) > a\} = \bigcup_{b>a} \{\ell_1(c) > b\}.$$

De lo anterior es claro que la función es decreciente y continua por la derecha.

### 3. Lema

*El conjunto  $\mathcal{Z}$  es casi seguramente acotado o no acotado acorde a si  $\bar{\Pi}(\infty)$  es positivo o nulo.*

Demostración: Por la propiedad fuerte de Markov aplicada en  $d_n(c)$ , vemos que  $\{\ell_n(c) : n \geq 1\}$  conforma una sucesión de variables aleatorias independientes e idénticamente distribuidas. Observemos que el evento

$$B = \{\text{existe un intervalo de excursión con } \ell > c \text{ tal que su longitud es infinita}\}$$

puede ser reescrito como

$$B = \bigcup_{n \geq 1} \left( \{\ell_n(c) = \infty\} \cap \bigcap_{i=1}^{n-1} \{\ell_i(c) < \infty\} \right).$$

Al poner  $p = \mathbb{P}(\ell_1(c) = \infty)$  notamos que la reescritura de B implica

$$\mathbb{P}(B) = \sum_{n \geq 1} p(1-p)^{n-1}.$$

Es claro que esta probabilidad es 1 si  $p > 0$  y 0 cuando  $p = 0$ . Resta notar que en realidad  $B = \{\mathcal{Z} \text{ es acotado}\}$  para concluir. ■

Acorde a lo anterior diremos que 0 es *transitorio* si  $\bar{\Pi}(\infty) > 0$  y *recurrente* si  $\bar{\Pi}(\infty) = 0$ .

### 4. Lema

*La función  $\bar{\Pi}$  cumple la propiedad*

$$(2) \quad \lim_{a \rightarrow 0^+} \bar{\Pi}(a) = \infty.$$

Demostración: Supongamos que (2) no se cumple, es decir que existe  $\varepsilon > 0$  tal que para todo  $\delta \in (0, c)$ ,  $\mathbb{P}(\ell_1(\delta) > c) \geq \varepsilon$ . De lo anterior se deduce que no hay intervalos de excursión en  $[0, g_1(c))$  con probabilidad positiva. Debido a que 0 es regular,  $\mathbb{P}(g_1(c) = 0) = 0$ . Luego, por la misma regularidad de 0 aunada a que X tiene trayectorias càd, la probabilidad de que X se quede en cero en una vecindad del origen es positiva. Esto viola el supuesto que hicimos sobre la instantaneidad del cero. ■

### 5. Lema

*Para cualesquiera  $a \in (0, \infty)$  y  $b \leq a$  con  $\bar{\Pi}(b) > 0$ ,*

$$\mathbb{P}(\ell_1(b) > a) = \frac{\bar{\Pi}(a)}{\bar{\Pi}(b)}.$$

Demostración: Para  $b = c$  o  $a = c$  simplemente estamos refraseando la definición:

$$\mathbb{P}(\ell_1(c) > a) = \bar{\Pi}(a) \quad \text{o} \quad \mathbb{P}(\ell_1(b) > c) = \frac{1}{\bar{\Pi}(b)}.$$

Veamos el resto de los casos, suponiendo primero que  $b < c < a$ . Bajo este supuesto,

$$\begin{aligned} \mathbb{P}(\ell_1(b) > a) &= \mathbb{P}(\ell_1(b) > c, \ell_1(c) > a) \\ &= \mathbb{P}(\ell_1(c) > a) - \mathbb{P}(\ell_1(b) \leq c, \ell_1(c) > a). \end{aligned}$$

Al aplicar la propiedad fuerte de Markov en  $d_1(b)$  obtenemos la independencia entre  $\{\ell_1(b) \leq c\}$  y  $\{\ell_1(c) > a\}$ , de modo que

$$\mathbb{P}(\ell_1(b) \leq c, \ell_1(c) > a) = \mathbb{P}(\ell_1(b) \leq c) \mathbb{P}(\ell_1(c) > a).$$

Entonces, por lo anterior y la definición de  $\bar{\Pi}$ ,

$$\mathbb{P}(\ell_1(b) > a) = \bar{\Pi}(a) - \left(1 - \frac{1}{\bar{\Pi}(b)}\right) \bar{\Pi}(a) = \frac{\bar{\Pi}(a)}{\bar{\Pi}(b)}.$$

Ahora supongamos que  $b < a < c$ . Mediante cuentas similares, aplicando la propiedad fuerte de Markov en  $d_1(b)$ , obtendremos el resultado. Efectivamente, tenemos

$$\begin{aligned} \mathbb{P}(\ell_1(b) > c) &= \mathbb{P}(\ell_1(b) > a, \ell_1(a) > c) \\ &= \mathbb{P}(\ell_1(a) > c) - \mathbb{P}(\ell_1(b) \leq a, \ell_1(a) > c), \end{aligned}$$

de donde

$$\frac{1}{\bar{\Pi}(b)} = \frac{1}{\bar{\Pi}(a)} \left(1 - (1 - \mathbb{P}(\ell_1(b) > a))\right) = \frac{1}{\bar{\Pi}(a)} \mathbb{P}(\ell_1(b) > a).$$

Esto nos da el resultado.

Finalmente consideramos el caso  $c < b < a$ . Ahora aplicaremos la propiedad fuerte de Markov en  $d_1(c)$  para deducir

$$\begin{aligned} \mathbb{P}(\ell_1(c) > a) &= \mathbb{P}(\ell_1(c) > b, \ell_1(b) > a) \\ &= \mathbb{P}(\ell_1(b) > a) - \mathbb{P}(\ell_1(c) \leq b) \mathbb{P}(\ell_1(b) > a) \\ &= \mathbb{P}(\ell_1(b) > a) \mathbb{P}(\ell_1(c) > b). \end{aligned}$$

Esto es equivalente a

$$\bar{\Pi}(a) = \bar{\Pi}(b) \mathbb{P}(\ell_1(b) > a),$$

siendo lo buscado. ■

Una consecuencia inmediata del lema 5 es que el cambiar la constante  $c$  modifica la función  $\bar{\Pi}$  por una constante multiplicativa.

Habiendo visto algunas de las propiedades de  $\ell_1(\cdot)$ , estudiaremos la cantidad de excursiones que inician antes de un punto dado. Para ello, dados  $a > 0$  y  $t > 0$  introducimos la notación

$$N_a(t) = \sup\{n \in \mathbb{N} : g_n(a) < t\}.$$

Por definición, esta variable cuenta el número de intervalos de excursión con  $\ell > a$  que comienzan antes de  $t$ , es decir que si  $X_t \neq 0$ , el intervalo de excursión al que pertenece  $t$  se cuenta si su longitud es mayor que  $a$ . En tanto, si  $t$  es el extremo izquierdo de un intervalo de excursión, tal intervalo se descarta. En el siguiente resultado establecemos la distribución de  $N_b$  cuando se evalúa en el extremo izquierdo de ciertos intervalos de excursión.

### 6. Proposición

Sean  $a \in (0, \infty]$  y  $b < a$  tales que  $\bar{\Pi}(b) > 0$ . Entonces  $N_b(g_1(a))$  es independiente de  $X \circ \theta_{g_1(a)}$  y tiene distribución geométrica de parámetro  $1 - \bar{\Pi}(a)/\bar{\Pi}(b)$ .

Demostración: Al usar la propiedad de Markov en el tiempo de paro  $d_n(b)$ , condicional en  $\{d_n(b) < \infty\}$ ,  $X \circ \theta_{d_n(b)}$  tiene la misma distribución que  $X$  y es independiente de  $\mathcal{F}_{d_n(b)}$ . Puesto que una excursión con longitud  $\ell > a$  también tiene longitud mayor que  $b$  deducimos

$$\{N_b(g_1(a)) \geq n\} = \{d_n(b) < g_1(a)\} = \{d_n(b) < d_1(a)\} \in \mathcal{F}_{d_n(b)},$$

que junto a lo previamente observado implica que para todo funcional medible  $G \geq 0$  se da la igualdad

$$\mathbb{E}[G(X \circ \theta_{d_n(b)}) \mathbb{1}_{\{N_b(g_1(a)) \geq n\}}] = \mathbb{E}[G(X)] \mathbb{P}(d_n(b) < g_1(a)).$$

En el evento  $\{d_n(b) < g_1(a)\}$ , el proceso  $X \circ \theta_{g_1(a)}$  se puede obtener de  $X^{(n)} = X \circ \theta_{d_n(b)}$  usando el operador  $\theta_{g_1^{(n)}(a)}$ , donde  $g_1^{(n)}(a)$  es el extremo izquierdo del primer intervalo de excursión  $X^{(n)}$  con longitud  $\ell > a$ . Por consiguiente obtenemos la independencia afirmada en el enunciado.

En cuanto a la distribución de  $N_b(g_1(a))$  vemos que

$$\begin{aligned} \mathbb{P}(N_b(g_1(a)) \geq n+1 \mid N_b(g_1(a)) \geq n) &= \mathbb{P}(d_{n+1}(b) < g_1(a) \mid d_n(b) < g_1(a)) \\ &= \mathbb{P}(d_1(b) < g_1(a)) \\ &= \mathbb{P}(\ell_1(b) \leq a) = 1 - \frac{\bar{\Pi}(a)}{\bar{\Pi}(b)}, \end{aligned}$$

donde la segunda igualdad se da por la propiedad fuerte de Markov aplicada en  $d_n(b)$ , mientras que la última se debe al lema 5. Esto muestra que  $N_b(g_1(a))$  tiene la distribución requerida. ■

### 7. Proposición

Para cada  $u \in (0, \infty]$  con  $\bar{\Pi}(u) > 0$ , el proceso

$$\frac{N_a(d_1(u))}{\bar{\Pi}(a)}, \quad a \in (0, u),$$

es una martingala retrógrada uniformemente integrable y continua por la derecha. Ella converge casi seguramente y en  $L^1(\mathbb{P})$  conforme  $a \rightarrow 0$ . El límite tiene distribución exponencial de parámetro  $\bar{\Pi}(u)$  y es independientes de  $\ell_1(u)$ .

Demostración: Para  $a < u$  consideremos la  $\sigma$ -álgebra generada por las longitudes de los intervalos de excursión con longitud  $\ell > a$  completadas antes de  $d_1(u)$ ; esto es

$$\mathcal{L}_a = \sigma(\ell_k(a), k \in \llbracket 1, N_a(d_1(u)) \rrbracket).$$

Es claro que  $N_a(d_1(u))$  es medible con respecto a  $\mathcal{L}_a$ . En adición, como toda excursión de longitud  $\ell > a$  es de longitud  $\ell > b$  siempre que  $b < a$ , tenemos que  $\mathcal{L}_a \subset \mathcal{L}_b$  cuando  $b < a$ , es decir que  $\{\mathcal{L}_a : a \in (0, u)\}$  es una filtración inversa.

Descomponemos la trayectoria de  $X$  en los intervalos de la forma  $[d_{k-1}(a), d_k(a)]$ , usando la convención  $d_0(a) = 0$ . De la propiedad fuerte de Markov en  $d_k(a)$ , condicional en el evento

$$\{N_a(d_1(u)) = n\} \cap \bigcap_{k=1}^n \{\ell_k(a) = \lambda_k\},$$

los procesos trasladados

$$Y^{(k)} := \{X_{t+d_{k-1}(a)} : 0 \leq t \leq d_k(a) - d_{k-1}(a)\}, \quad k \in \llbracket 1, n \rrbracket,$$

son independientes y cada  $Y^{(k)}$  tiene la misma ley que el proceso  $\{X_t : 0 \leq t \leq d_1(a)\}$  condicional en  $\ell_1(a) = \lambda_k$ . Sabemos, en vista de la demostración realizada de la proposición 6, que para  $b < a$ ,

$$N_b(d_1(a)) = N_b(g_1(a)) + 1.$$

Por la misma proposición 6,  $N_b(g_1(a))$  es independiente de  $\ell_1(a)$  y tiene distribución geométrica  $1 - \bar{\Pi}(a)/\bar{\Pi}(b)$ . Consecuentemente, la distribución de  $N_b(d_1(u))$  condicional en  $\mathcal{L}_a$  es la misma que la de la suma aleatoria

$$(\xi_1 + 1) + \cdots + (\xi_n + 1),$$

donde  $n = N_a(d_1(u))$  y  $\{\xi_k : k \geq 1\}$  es una sucesión de variables aleatorias independientes con misma distribución geométrica de parámetro  $1 - \bar{\Pi}(a)/\bar{\Pi}(b)$ . Puesto que

$$\mathbb{E}[\xi_1] = \frac{1 - \bar{\Pi}(a)/\bar{\Pi}(b)}{\bar{\Pi}(a)/\bar{\Pi}(b)},$$

obtenemos  $\mathbb{E}[\xi_1 + 1] = \bar{\Pi}(b)/\bar{\Pi}(a)$ . Por consiguiente,

$$\mathbb{E}[N_b(d_1(u)) \mid \mathcal{L}_a] = N_a(d_1(u)) \frac{\bar{\Pi}(b)}{\bar{\Pi}(a)},$$

con lo cual probamos la propiedad de martingala retrógrada. La continuidad por la derecha se sigue por construcción. En adición, es posible construir una versión de la martingala que sea continua por la izquierda.

El hecho de que sea una martingala retrógrada positiva implica que ella converge casi seguramente conforme  $a \rightarrow 0$ . Ahora bien, debido a que la proposición 6 nos dice que la variable aleatoria  $N_a(d_1(u)) = N_a(g_1(u)) + 1$  es independiente del proceso  $X \circ \theta_{g_1(u)}$ , vemos que ella también es independiente de  $\ell_{N_a(d_1(u))}(a) = \ell_1(u)$ . Así,  $\ell_1(u)$  es independiente de

$$\left\{ \frac{N_a(d_1(u))}{\bar{\Pi}(a)} : 0 < a < u \right\},$$

lo cual implica que  $\ell_1(u)$  es también independiente del límite.

Resta verificar que el límite tiene distribución exponencial, además de la convergencia en  $L^1(\mathbb{P})$ . Para lo primero, notamos que  $N_a(d_1(u)) - 1$  tiene distribución

geométrica de parámetro  $1 - \bar{\Pi}(u)/\bar{\Pi}(a)$ . En consecuencia, para  $\lambda > 0$ ,

$$\begin{aligned} \mathbb{E} \left[ \exp \left\{ -\lambda \frac{N_a(d_1(u))}{\bar{\Pi}(a)} \right\} \right] &= e^{-\lambda/\bar{\Pi}(a)} \frac{\bar{\Pi}(u)/\bar{\Pi}(a)}{1 - e^{-\lambda/\bar{\Pi}(a)}(1 - \bar{\Pi}(u)/\bar{\Pi}(a))} \\ &= e^{-\lambda/\bar{\Pi}(a)} \frac{\bar{\Pi}(u)}{e^{-\lambda/\bar{\Pi}(a)}\bar{\Pi}(u) - \bar{\Pi}(a)(e^{-\lambda/\bar{\Pi}(a)} - 1)} \\ &\xrightarrow{a \rightarrow 0} \frac{\bar{\Pi}(u)}{\bar{\Pi}(u) + \lambda}, \end{aligned}$$

donde la convergencia se da debido al lema 4. Esto muestra que

$$\lim_{a \rightarrow 0} \frac{N_a(d_1(u))}{\bar{\Pi}(a)}$$

tiene la distribución exponencial adecuada. Cuentas análogas revelan la convergencia

$$\lim_{a \rightarrow 0} \mathbb{E} \left[ \frac{N_a(d_1(u))}{\bar{\Pi}(a)} \right] = \frac{1}{\bar{\Pi}(u)}.$$

Por el lema de Scheffé esto implica la convergencia en  $L^1(\mathbb{P})$  de la martingala, lo cual nos da la integrabilidad uniforme. ■

Lo realizado nos permitirá construir el tiempo local.

### 8. Teorema

*Las siguientes afirmaciones suceden casi seguramente:*

- (i) *Para toda  $t \geq 0$ ,  $N_a(t)/\bar{\Pi}(a)$  converge conforme  $a \rightarrow 0$ . Denotamos el límite por  $L_t$ .*
- (ii) *El mapeo  $t \mapsto L_t$  es creciente y continuo. Él es llamado el tiempo local de  $X$  en 0.*
- (iii) *El soporte de la medida de Stieltjes  $dL$  es  $\text{cl}(\mathcal{X})$ .*

**Demostración:** Comenzamos demostrando los puntos (i) y (ii). Gracias a la proposición 7 sabemos que  $N_a(t)/\bar{\Pi}(a)$  converge a  $L_t$  cuando  $t = d_1(u)$ . Por la propiedad de Markov aplicada de manera iterada en los tiempos de paro  $d_k(u)$ ,  $k \in \mathbb{N}$ , y tomando la unión sobre  $u > 0$ , la convergencia se da para todo  $t \in \mathcal{D}$ , donde

$$\mathcal{D} = \{d_k(u) : u > 0, k \in \mathbb{N}\}$$

es el conjunto de extremos derechos de los intervalos de excursión.

El mapeo  $L : \mathcal{D} \rightarrow [0, \infty)$  es claramente creciente. Probaremos que tiene una única extensión creciente a  $[0, \infty)$ . En este sentido fijemos  $\varepsilon > 0$  y consideremos el evento

$$\Lambda_a = \{L_{d_k(a)} - L_{d_{k-1}(a)} \leq \varepsilon \text{ para toda } k \leq N_a(d_1(c))\}.$$

Sabemos, por la proposición 6, que  $N_a(d_1(c)) - 1$  tiene distribución geométrica de parámetro  $1 - \bar{\Pi}(c)/\bar{\Pi}(a)$ . También sabemos, por la proposición 7 y la propiedad fuerte de Markov, que condicional en  $N_a(d_1(c)) = n$ , las variables aleatorias  $L_{d_k(a)} - L_{d_{k-1}(a)}$ ,  $k \in \llbracket 1, n \rrbracket$ , son independientes y tienen distribución exponencial de parámetro  $\bar{\Pi}(a)$ .

Esto nos permite calcular la probabilidad de  $\Lambda_a$ ,

$$\begin{aligned}
 \mathbb{P}(\Lambda_a) &= \mathbb{E} \left[ \prod_{k=1}^{N_a(d_1(c))} \mathbb{1}_{\{L_{d_k(a)} - L_{d_{k-1}(a)} \leq \varepsilon\}} \right] \\
 &= \sum_{n \geq 1} \mathbb{P}(N_a(d_1(c)) = n) \prod_{k=1}^n \mathbb{P}(L_{d_k(a)} - L_{d_{k-1}(a)} \leq \varepsilon) \\
 &= \sum_{n \geq 1} \frac{\bar{\Pi}(c)}{\bar{\Pi}(a)} \left(1 - \frac{\bar{\Pi}(c)}{\bar{\Pi}(a)}\right)^{n-1} (1 - e^{-\varepsilon \bar{\Pi}(a)})^n \\
 &= \bar{\Pi}(c) (1 - e^{-\varepsilon \bar{\Pi}(a)}) \left[ \bar{\Pi}(a) e^{-\varepsilon \bar{\Pi}(a)} + \bar{\Pi}(c) (1 - e^{-\varepsilon \bar{\Pi}(a)}) \right]^{-1}.
 \end{aligned}$$

Notando que si  $a < b$ ,  $\Lambda_a \supset \Lambda_b$ , del cálculo anterior y el lema 4 se deduce

$$\mathbb{P}\left(\bigcup_{a < c} \Lambda_a\right) = 1.$$

Esto quiere decir que, con probabilidad uno, no existe una brecha mayor que  $\varepsilon$  en  $L(\mathcal{D} \cap [0, d_1(c)])$ . Considerando  $\varepsilon \rightarrow 0^+$  y usando la propiedad fuerte de Markov, de forma iterada, en los tiempos  $d_k(c)$ , obtenemos que  $L(\mathcal{D})$  es denso en  $[L_0, L_\infty)$  con probabilidad uno. Por ende, casi seguramente,

$$(3) \quad \inf\{L_t : t \in \mathcal{D}, t > s\} = \sup\{L_t : t \in \mathcal{D}, t < s\}, \quad \text{para toda } s \in [0, \infty).$$

Al definir a  $L_s$  por (3), vemos que  $L : [0, \infty) \rightarrow [0, \infty)$  es continua, además de ser la única extensión creciente de  $L : \mathcal{D} \rightarrow [0, \infty)$ .

Dado  $a > 0$ , es claro que

$$\sup\{N_a(t) : t \in \mathcal{D}, t < s\} \leq N_a(s) \leq \inf\{N_a(t) : t \in \mathcal{D}, t > s\}.$$

Entonces, por la proposición 7, la expresión (3) y la continuidad de  $L$ , concluimos que con probabilidad uno,  $N_a(s)/\bar{\Pi}(a)$  converge a  $L_s$  para toda  $s \in [0, \infty)$ .

Por la construcción realizada para verificar (i), tenemos  $\text{supp } dL \subset \text{cl}(\mathcal{Z})$ . Aún debemos demostrar que, con probabilidad uno, para cualesquiera  $0 < s < t$ ,  $L_s < L_t$  siempre que  $X$  visite 0 en el intervalo abierto  $(s, t)$ . Sin pérdida de generalidad, gracias a la monotonía de  $L$ , podemos considerar  $s$  y  $t$  en un conjunto contable denso y entonces será suficiente probar que para  $0 < s < t$  dados,

$$(4) \quad \mathbb{P}(L_s = L_t \text{ y } X_u = 0 \text{ para algún } u \in (s, t)) = 0.$$

Dada  $\nu > 0$ , la regularidad e instantaneidad de 0 implican que para cualquier  $\varepsilon > 0$  existe  $a > 0$  tal que  $\mathbb{P}(d_1(a) < \nu) > 1 - \varepsilon$ . Debido a la proposición 7 sabemos que  $\mathbb{P}(L_{d_1(a)} = 0) = 0$  y en consecuencia, usando que  $L$  es creciente,

$$\mathbb{P}(L_\nu = 0) \leq \mathbb{P}(d_1(a) \geq \nu) < \varepsilon.$$

Así, hemos demostrado que  $\mathbb{P}(L_\nu = 0) = 0$  para toda  $\nu > 0$ , lo cual basta para establecer (4) a través del uso de la propiedad fuerte de Markov en el tiempo de paro  $R_s$  pues (4) se puede reescribir como

$$\mathbb{P}(L_s = L_t, R_s < t) = \int_s^t \mathbb{P}(L_{t-r} = 0) \mathbb{P}(R_s \in dr) = 0.$$

Con este argumento, se ha demostrado que  $\text{cl}(\mathcal{Z}) \subset \text{supp } dL$ , lo cual concluye la demostración. ■

La construcción realizada de  $L$  depende de la constante  $c$ . Sin embargo, como la constante  $c$  únicamente afecta a  $\bar{\Pi}$  por un factor multiplicativo, lo mismo sucede con  $L$ .

Al usar la propiedad fuerte de Markov en la primera llegada al origen, podemos extender la construcción del tiempo local de  $X$  al caso en el cual  $X$  no inicia en 0. Con esta idea, para cualquier tiempo de paro  $T < \infty$ , al denotar por  $L' = \{L'_t : t \geq 0\}$  el tiempo local de 0 del proceso  $X \circ \theta_T$ , por el punto (i) del teorema 8 tendremos que

$$(5) \quad L_{T+t} = L_T + L'_t \quad \text{para todo } t \geq 0 \text{ casi seguramente.}$$

Decimos que  $L$  es un *funcional aditivo*, ver la definición 1.1 del capítulo iv de Blumenthal y Gettoor (1968). Más aún, la propiedad fuerte de Markov implica que para todo tiempo de paro con la propiedad  $X_T = 0$  casi seguramente en el evento  $\{T < \infty\}$ , bajo la medida de probabilidad  $\mathbb{P}(\cdot \mid T < \infty)$ , el proceso

$$\{(X \circ \theta_T, L_{T+t} - L_T) : t \geq 0\}$$

es independiente de  $\mathcal{F}_T$  y tiene la misma ley que  $(X, L)$ .

De la misma construcción hacemos notar que el proceso  $L$  es adaptado a la filtración  $\{\mathcal{F}_t : t \geq 0\}$ . Cerraremos la sección viendo que el proceso  $L$  construido en el teorema 8 es el único proceso continuo y adaptado que crece únicamente en  $\text{cl}(\mathcal{Z})$  y tiene la propiedad aditiva (5).

### 9. Proposición

Sea  $A = \{A_t : t \geq 0\}$  un proceso continuo y creciente adaptado a  $\{\mathcal{F}_t : t \geq 0\}$  que satisface las propiedades:

- (i) El soporte de la medida de Stieltjes  $dA$  está contenido, casi seguramente, en  $\text{cl}(\mathcal{Z})$ .
- (ii) Para todo tiempo de paro  $T$  para el cual  $X_T = 0$  casi seguramente en el evento  $\{T < \infty\}$ , el proceso trasladado

$$\{(X \circ \theta_T, A_{T+t} - A_T) : t \geq 0\}$$

es independiente de  $\mathcal{F}_T$  bajo  $\mathbb{P}(\cdot \mid T < \infty)$ . En adición, el proceso trasladado tiene la misma ley que  $(X, A)$  bajo  $\mathbb{P}$ .

Entonces existe una constante  $k \geq 0$  tal que  $A = kL$  con probabilidad uno.

*Demostración:* Dado  $b > 0$ , afirmamos que  $A_{d_1(b)}$  tiene distribución exponencial. Para verificar la afirmación primero fijamos  $s, t > 0$  y definimos  $T = \inf\{u : A_u > t\}$ . En vista de que  $A$  es continuo y creciente  $\{T < \nu\} = \{A_\nu > t\}$  para toda  $\nu$ , lo cual implica que  $T$  es un tiempo de paro debido a que  $A$  es adaptado y la filtración es càd. Procedemos a trabajar condicionalmente en el evento  $\{T < \infty\}$ . Puesto que  $A$  es continuo,  $A_T = t$  y entonces  $X_T = 0$  ya que  $A$  solo crece en  $\text{cl}(\mathcal{Z})$ . En efecto, si  $X_T \neq 0$ , por la continuidad derecha de las trayectorias de  $X$ ,  $X_{T+r} \neq 0$  para  $r$  suficientemente pequeño, de modo que  $A_{T+r} = t$  y en consecuencia  $T < \inf\{u : A_u > t\}$ , siendo esto un sinsentido. Luego, por la condición (ii) el proceso  $(X \circ \theta_T, A')$ , con  $A' = A \circ \theta_T - t$ , tiene la misma ley que  $(X, A)$  bajo  $\mathbb{P}$ . Si denotamos por  $d'_1(b)$  al extremo derecho de la primera excursión de  $X \circ \theta_T$  con  $\ell > b$ , entonces

$$\{A_{d_1(b)} > t, A'_{d'_1(b)} > s\} = \{A_{d_1(b)} > A_T, A_{T+d'_1(b)} > t + s\} = \{A_{d_1(b)} > t + s\}.$$

Con esta igualdad y la condición (ii) deducimos

$$\begin{aligned}\mathbb{P}(A_{d_1(b)} > t + s) &= \mathbb{P}(A_{d_1(b)} > t) \mathbb{P}(A'_{d'_1(b)} > s \mid T < d_1(b)) \\ &= \mathbb{P}(A_{d_1(b)} > t) \mathbb{P}(A_{d_1(b)} > s),\end{aligned}$$

mostrando que  $A_{d_1(b)}$  tiene distribución exponencial. Denotemos por  $\lambda(b)$  su parámetro. Podemos suponer, sin pérdida de generalidad, que  $\lambda(b) < \infty$  para toda  $b \in (0, c]$ , pues en caso contrario,  $A$  sería idénticamente cero y no habría nada que probar.

Con el supuesto anterior, verificamos que  $\lambda$  y  $\bar{\Pi}$  son proporcionales. Usando la propiedad fuerte de Markov en el tiempo  $T_v = \inf\{u : A_u > v\}$ , vemos que

$$\mathbb{P}(A_{d_1(b)} > v, \ell_1(b) > s) = \mathbb{P}(T_v < d_1(b), \ell_1(b) > s) = \mathbb{P}(A_{d_1(b)} > v) \mathbb{P}(\ell_1(b) > s),$$

es decir que  $A_{d_1(b)}$  es independiente de  $\ell_1(b)$ . Visto lo anterior, fijemos  $a > b$ . Al aplicar la condición (ii) en los tiempos  $d_k(b)$ , con  $k \in \llbracket 1, N_b(g_1(a)) \rrbracket$ , deducimos

$$A_{d_1(a)} = \xi_1 + \cdots + \xi_n, \quad \text{donde } n = N_b(g_1(a)) + 1.$$

Las variables aleatorias  $\{\xi_i : i \geq 1\}$  independientes y tienen distribución exponencial de parámetro  $\lambda(b)$ . Por el argumento realizado antes, estas variables aleatorias también son independientes de  $N_b(g_1(a))$ . Sabemos, de la proposición 6, que  $N_b(g_1(a))$  tiene distribución geométrica de parámetro  $1 - \bar{\Pi}(a)/\bar{\Pi}(b)$ . Por lo tanto,

$$\frac{1}{\lambda(a)} = \mathbb{E}[A_{d_1(a)}] = \mathbb{E}\left[\sum_{i=1}^{N_b(g_1(a))+1} \xi_i\right] = \frac{1}{\lambda(b)} \mathbb{E}[N_b(g_1(a)) + 1] = \frac{1}{\lambda(b)} \frac{\bar{\Pi}(b)}{\bar{\Pi}(a)}.$$

En consecuencia existe una constante  $k > 0$  tal que  $k\lambda = \bar{\Pi}$ .

Para concluir, notemos que

$$\begin{aligned}\mathbb{E}\left[\left|A_{d_1(a)} - \frac{N_b(d_1(a))}{\lambda(b)}\right|^2\right] &= \frac{1}{\lambda(b)^2} \mathbb{E}\left[\left|\sum_{i=1}^{N_b(d_1(a))} (\lambda(b)\xi_i - 1)\right|^2\right] \\ &= \frac{1}{\lambda(b)^2} \text{Var}(\lambda(b)\xi_1) \mathbb{E}[N_b(d_1(a))] \\ &= \frac{1}{\lambda(b)^2} \frac{\bar{\Pi}(b)}{\bar{\Pi}(a)} = \frac{k^2}{\bar{\Pi}(b)\bar{\Pi}(a)} \xrightarrow{b \rightarrow 0} 0.\end{aligned}$$

Usando la misma técnica con la que se probó la parte (i) del teorema 8, la convergencia previa nos permite deducir  $A_t = kL_t$  para toda  $t \in \mathcal{D}$ . En vista de que  $A$  y  $L$  únicamente crecen en  $\text{cl}(\mathcal{Z})$ , la identidad se extiende a  $\text{cl}(\mathcal{Z})^c$ , conjunto que es denso en  $\mathbb{R}_+$  gracias a la proposición 1. Por tanto, concluimos  $A = kL$  debido a la continuidad de ambos procesos. ■

## 10. Corolario

Existe una constante  $\mathfrak{d}$  tal que, con probabilidad uno,

$$\int_0^t \mathbb{1}_{\{X_s=0\}} ds = \int_0^t \mathbb{1}_{\{s \in \text{cl}(\mathcal{Z})\}} ds = \mathfrak{d}L_t \quad \text{para toda } t \geq 0.$$

Demostración: El conjunto  $\text{cl}(\mathcal{Z})$  difiere de  $\mathcal{Z}$  por a lo más una cantidad contable de puntos pues  $X$  tiene a lo más una cantidad contable de discontinuidades. Lo anterior nos dice que ambas integrales coinciden. Es claro que las integrales satisfacen la condición (i) de la proposición 9, mientras que la condición (ii) es una consecuencia de la propiedad fuerte de Markov de  $X$  y la aditividad de las integrales. ■

## 2.3. Tiempo local inverso

El objetivo de la presente sección es describir la distribución del tiempo local  $L$  caracterizando la ley de su inversa continua por la derecha:

$$L_t^{-1} = \inf\{s \geq 0 : L_s > t\}, \quad t \geq 0.$$

En lo subsiguiente también será útil tener notación para la inversa continua por la izquierda:

$$L_{t-}^{-1} = \inf\{s \geq 0 : L_s \geq t\} = \lim_{s \rightarrow t-} L_s, \quad t \geq 0.$$

Iniciamos estudiando algunas de las propiedades elementales del tiempo local inverso.

### 1. Proposición

- (i) Para cada  $t \geq 0$ , tanto  $L_t^{-1}$  como  $L_{t-}^{-1}$  son tiempos de paro.
- (ii) El proceso  $L^{-1}$  es creciente, continuo por la derecha y adaptado a la filtración  $\{\mathcal{F}_{L_t^{-1}} : t \geq 0\}$ .
- (iii) Casi seguramente, para toda  $t > 0$ ,

$$L_{L_t^{-1}}^{-1} = \inf\{L_u^{-1} \geq 0 : L_u^{-1} > t\} = \inf\{s > t : X_s = 0\}$$

$\gamma$

$$L_{L_{t-}^{-1}}^{-1} = \sup\{L_u^{-1} \geq 0 : L_u^{-1} < t\} = \sup\{s < t : X_s = 0\}.$$

En particular,  $L_t^{-1} \in \mathcal{Z}$  en el evento  $\{L_t^{-1} < \infty\}$ .

Demostración: Por la continuidad de  $L$ , para  $s, t > 0$  se tiene que

$$\{L_t^{-1} < s\} = \{L_s > t\} \in \mathcal{F}_s.$$

Entonces  $L_t^{-1}$  es un tiempo de paro debido a que  $\{\mathcal{F}_t : t \geq 0\}$  es càd. Que  $L_{t-}^{-1}$  es un tiempo de paro se seguirá al verificar que  $t \mapsto L_t^{-1}$  es creciente pues en ese caso  $L_{t-}^{-1}$  será el límite creciente de tiempos de paro, y por lo tanto un tiempo de paro.

Por definición,  $t \mapsto L_t^{-1}$  es creciente y continua por la derecha. En efecto, por ser  $L$  creciente esta afirmación se sigue de la igualdad de conjuntos

$$\{s \geq 0 : L_s > t\} = \bigcup_{r > t} \{s \geq 0 : L_s > r\}.$$

El hecho de que  $L$  sea adaptado a la filtración  $\{\mathcal{F}_{L_t^{-1}} : t \geq 0\}$  se sigue del numeral (i).

Demostremos ahora que se cumple el punto (iii). Primero notemos que se tienen las contenciones

$$\{u > L_t\} \subset \{L_u^{-1} > t\} \subset \{u \geq L_t\}.$$

Entonces, del punto (ii) y estas contenciones obtenemos

$$\begin{aligned} L_{L_t^{-1}}^{-1} &= \inf\{L_u^{-1} \geq 0 : u > L_t\} \\ &\geq \inf\{L_u^{-1} \geq 0 : L_u^{-1} > t\} \\ &\geq \inf\{L_u^{-1} \geq 0 : u \geq L_t\} = L_t^{-1}. \end{aligned}$$

Esto prueba la primera igualdad del punto (iii). Consideremos ahora el tiempo de primer retorno al cero después del tiempo  $t$ ,

$$R_t = \inf\{s > t : X_s = 0\}.$$

Probaremos que, con probabilidad uno  $R_t = L_{L_t}^{-1}$ .

Supongamos primero que  $R_t > t$ . Debido al teorema 2.8,  $L$  es constante en el intervalo  $[t, R_t)$  y por lo tanto  $R_t \leq L_{L_t}^{-1}$ . En el caso  $R_t = \infty$  no hay nada que probar, de modo que supondremos que  $R_t < \infty$ . Bajo este supuesto,  $R_t \in \text{supp } dL$  es un punto aislado por la izquierda. Por otra parte, por ser  $L$  continuo,  $R_t$  no puede ser aislado por la derecha. Por consiguiente,  $L_s > L_{R_t} = L_t$  para toda  $s > R_t$ , mediante lo cual deducimos  $R_t \geq L_{L_t}^{-1}$ .

Ahora supongamos que  $R_t = t$ . Entonces  $t \in \text{supp } dL$  no es aislado por la derecha. Argumentando de manera similar a como se hizo antes,  $L_s > L_t$  para toda  $s > t$ , y en consecuencia  $t \geq L_{L_t}^{-1}$ . La desigualdad inversa es evidente de la definición de  $L^{-1}$ . Esto prueba

$$L_{L_t}^{-1} = \inf\{L_u^{-1} \geq 0 : L_u^{-1} > t\} = R_t.$$

En cuanto al segundo conjunto de igualdades en (iii), ellas se siguen de argumentos análogos a los ya expuestos. Para concluir notemos que en  $\{L_t^{-1} < \infty\}$  existe  $s \geq 0$  tal que  $L_s = t$ . La primera identidad implica  $L_t^{-1} \in \mathcal{Z}$ . ■

Antes de caracterizar la ley de  $L^{-1}$ , hacemos notar que el hecho de que  $L$  sea continua junto a que  $\text{supp } dL = \text{cl}(\mathcal{Z})$  nos dicen que  $\text{cl}(\mathcal{Z})$  no tiene puntos aislados, situación que también conocíamos por la proposición 2.1. De esto y la propiedad (iii) de la proposición 1 deducimos que los intervalos de excursión son aquellos de la forma  $(L_t^{-1}, L_{t+}^{-1})$  cuando  $L_t^{-1} < L_{t+}^{-1}$ .

Para el siguiente resultado introducimos la medida  $\Pi$  sobre  $(0, \infty)$  mediante

$$\Pi((s, t]) = \bar{\Pi}(s) - \bar{\Pi}(t), \quad 0 < s < t < \infty.$$

## 2. Teorema

*El tiempo local inverso  $L^{-1}$  es un subordinador de Lévy con medida  $\Pi$ , coeficiente de deriva  $\mathfrak{d}$ , matado a tasa  $\bar{\Pi}(\infty)$ . Luego, para cualesquiera  $t, \lambda > 0$ ,*

$$\mathbb{E}[\exp\{-\lambda L_t^{-1}\}] = \exp\{-t\Phi(\lambda)\},$$

*donde el exponente de Laplace  $\Phi$  es dado por*

$$\Phi(\lambda) = \lambda\left(\mathfrak{d} + \int_0^\infty e^{-\lambda r} \bar{\Pi}(r) dr\right).$$

*Demostración:* Separaremos los casos  $\bar{\Pi}(\infty) = 0$  y  $\bar{\Pi}(\infty) > 0$  por claridad.

Supongamos pues que  $\bar{\Pi}(\infty) = 0$ , que por el lema 2.3 y la probabilidad en (2.1) nos dicen que con probabilidad uno  $d_1(c) < \infty$ . Aplicando la propiedad fuerte de Markov deducimos entonces que  $d_n(c) < \infty$  con probabilidad uno para cada  $n \in \mathbb{N}$ . Entonces, por la propiedad de Markov, la proposición 2.7 y la propiedad aditiva (2.5) de  $L$ ,  $L_{d_n(c)}$  se puede expresar como la suma de  $n$  variables aleatorias independientes con distribución exponencial de parámetro 1. Consecuentemente,  $L_\infty = \lim_n L_{d_n(c)} = \infty$  casi seguramente.

En vista del punto (i) de la proposición 1 podemos aplicar la propiedad fuerte de Markov en el tiempo  $L_t^{-1}$ , de tal forma que  $\tilde{X} = X \circ \theta_{L_t^{-1}}$  tiene la misma ley que  $X$  y es independiente de  $\mathcal{F}_{L_t^{-1}}$ . Al denotar por  $\tilde{L}$  al tiempo local de  $\tilde{X}$  y observar que  $L_{L_t^{-1}} = t$ , deducimos de (2.5), la igualdad

$$\tilde{L}_s = L_{L_t^{-1}+s} - t,$$

de modo que el tiempo local inverso de  $\tilde{X}$  es

$$(1) \quad \begin{aligned} \tilde{L}_s^{-1} &= \inf\{u \geq 0 : \tilde{L}_u > s\} = \inf\{u \geq 0 : L_{L_t^{-1}+u} > s + t\} \\ &= \inf\{u \geq 0 : L_t^{-1} + u > L_{s+t}^{-1}\} = L_{s+t}^{-1} - L_t^{-1}. \end{aligned}$$

Esto demuestra que  $L^{-1}$  tiene incrementos independientes y estacionarios. Como en adición es creciente y càd, concluimos que  $L^{-1}$  es un subordinador.

Recordemos que la medida de Lévy  $\nu$  de  $L^{-1}$  es la medida característica del proceso Poisson puntual de los saltos  $\Delta L^{-1}$ . Para cada  $a > 0$  denotamos por

$$T_a = \inf\{t \geq 0 : \Delta L_t^{-1} > a\}$$

al instante en el cual el primer salto tiene longitud  $\ell > a$ . De las observaciones realizadas posterior a la proposición 1,  $T_a$  coincide con el tiempo local  $L$  evaluado en el extremo derecho del primer intervalo de longitud  $\ell > a$ . Por consiguiente, debido a la proposición 2.7,  $T_a$  tiene distribución exponencial de parámetro  $\bar{\Pi}(a)$ . Luego, por propiedades de las medidas aleatorias de Poisson,  $\nu((a, \infty)) = \bar{\Pi}(a) = \Pi((a, \infty))$ , lo cual prueba que la medida de Lévy de  $L^{-1}$  es  $\Pi$ .

Ahora, usando la correspondencia entre  $\Delta L^{-1}$  y las longitudes de los intervalos de excursión, aunado que estos intervalos son aquellos que aparecen en la descomposición canónica del conjunto  $\mathbb{R}_+ \setminus \text{cl}(\mathcal{Z})$  obtenemos

$$\int_0^{L_t^{-1}} \mathbb{1}_{\{s \notin \text{cl}(\mathcal{Z})\}} ds = \sum_{s \leq t} \Delta L_s^{-1}.$$

Mientras tanto, gracias al corolario 2.10 y a que  $L_{L_t^{-1}} = t$ ,

$$\int_0^{L_t^{-1}} \mathbb{1}_{\{s \in \text{cl}(\mathcal{Z})\}} ds = \partial L_{L_t^{-1}} = \partial t.$$

Por ende,

$$L_t^{-1} = \partial t + \int_{s \leq t} \Delta L_s^{-1}.$$

Esto prueba que la deriva es precisamente  $\partial$ , con lo cual concluimos el caso  $\bar{\Pi}(\infty) = 0$ .

Supongamos ahora que  $\bar{\Pi}(\infty) > 0$ . Usando argumentos análogos a los del caso  $\bar{\Pi}(\infty) = 0$ , la propiedad fuerte de Markov y la aditividad de  $L$  implican que para cualesquiera  $0 < t < t'$ , condicional en  $\{L_{t'}^{-1} < \infty\}$ , la ley de  $\{L_s^{-1} : 0 \leq s \leq t\}$  coincide con la de un subordinador  $S$  restringido al intervalo  $[0, t]$ . Debido a la igualdad de conjuntos  $\{L_t^{-1} < \infty\} = \{L_\infty > t\}$ , la afirmación anterior se puede reexpresar como:  $\{L_s^{-1} : 0 \leq s < L_\infty\}$  tiene la misma distribución que  $\{S_s : 0 \leq s < \tau\}$ , donde  $\tau$  es una variable aleatoria independiente de  $S$  cuya distribución coincide con la de  $L_\infty$ . La distribución en cuestión es exponencial de parámetro  $\bar{\Pi}(\infty)$ , observación que se sigue de la proposición 2.7.

Entonces, si denotamos por  $\nu$  a la medida de Lévy de  $S$  y consideramos  $T_a$  como antes, vemos que para cada  $a > 0$ ,

$$\begin{aligned} 1 - \exp\{-t\nu((a, \infty))\} &= \mathbb{P}(\text{existe } s < t \text{ tal que } \Delta S_s > a) \\ &= \mathbb{P}(T_a < t \mid L_\infty > t) \\ &= \exp\{t\bar{\Pi}(\infty)\}\mathbb{P}(L_{d_1(a)} < t, L_\infty > t). \end{aligned}$$

Notemos que  $L_{d_1(a)}$ , condicional en  $\{d_1(a) < \infty\}$ , tiene distribución exponencial de parámetro  $\bar{\Pi}(a)$  en vista de la proposición 2.7. Más aún, por el lema 2.5,

$$\mathbb{P}(d_1(a) < \infty) = \mathbb{P}(\ell_1(a) < \infty) = 1 - \frac{\bar{\Pi}(\infty)}{\bar{\Pi}(a)}.$$

Subsecuentemente, aplicando la propiedad fuerte de Markov en  $d_1(a)$ ,

$$\begin{aligned} \mathbb{P}(L_{d_1(a)} < t, L_\infty > t) &= \mathbb{E}[\mathbb{1}_{\{L_{d_1(a)} < t\}} \mathbb{1}_{\{L_\infty > t\}} \mathbb{1}_{\{d_1(a) < \infty\}}] \\ &= \mathbb{E}\left[\int_0^t \bar{\Pi}(a) e^{-s\bar{\Pi}(a)} \mathbb{P}(L'_\infty > t-s) ds \mathbb{1}_{\{d_1(a) < \infty\}}\right] \\ &= \left(1 - \frac{\bar{\Pi}(\infty)}{\bar{\Pi}(a)}\right) \int_0^t \bar{\Pi}(a) e^{-(\bar{\Pi}(a)-\bar{\Pi}(\infty))s} e^{-\bar{\Pi}(\infty)t} ds \\ &= e^{-t\bar{\Pi}(\infty)} (1 - e^{-(\bar{\Pi}(a)-\bar{\Pi}(\infty))t}). \end{aligned}$$

De lo anterior vemos que  $\nu((a, \infty)) = \bar{\Pi}(a) - \bar{\Pi}(\infty) = \Pi((a, \infty))$ , mostrando que  $\nu = \Pi$ . Que la deriva es  $\partial$  se obtiene igual que en el caso  $\bar{\Pi}(\infty) = 0$ .

Para concluir, resta notar que la identidad del exponente de Laplace se sigue de la fórmula de Lévy–Kintchine e integración por partes. Alternativamente podemos usar el teorema de Tonelli para verificar

$$\begin{aligned} \int_0^\infty e^{-\lambda r} \bar{\Pi}(r) dr &= \int_0^\infty \int_{(r, \infty)} e^{-\lambda r} d\Pi(u) dr \\ &= \frac{1}{\lambda} \int_{(0, \infty)} (1 - e^{-\lambda u}) d\Pi(u). \end{aligned}$$

Concluimos por la fórmula de Lévy–Kintchine. ■

Una consecuencia de la caracterización hecha es que  $\Pi$  satisface la condición de integrabilidad

$$\int_{(0, \infty)} 1 \wedge u d\Pi(u) < \infty,$$

ya que los subordinadores son de variación finita.

## 2.4. Proceso de excursiones

Habiendo desarrollado teoría para estudiar las longitudes de los intervalos de excursión, ahora estamos interesados en las excursiones de  $X$  fuera de cero, es decir partes de la trayectoria de la forma  $\{X_{g+t} : 0 \leq t < d-g\}$  correspondientes a los intervalos de excursión  $(g, d)$ . Sabemos que los tiempos de vida de las excursiones coinciden con las longitudes de los intervalos de excursión. Antes de introducir una medida

$\sigma$ -finita que describe la distribución de las excursiones de  $X$ , haremos énfasis sobre el espacio trayectorial con el que se trabajará.

En lo que sigue consideraremos el espacio de trayectorias càdlàg  $\omega$  en

$$\Omega' = \mathbb{D}(\mathbb{R}_+, E \cup \{\Delta\}),$$

con tiempo de vida  $\zeta = \inf\{t \geq 0 : \omega(t) = \Delta\}$ . Este espacio lo consideramos con la topología  $J_1$  de Skorokhod, bajo la cual es un espacio polaco, ver la sección 16 de Billingsley (1999), sección III.5 de Ethier y Kurtz (2005) o la sección 2 de Whitt (1980).

Para describir el proceso de excursiones, introducimos el espacio de excursiones. Dado  $a > 0$ , denotemos por  $\mathcal{E}^{(a)}$  el conjunto de excursiones con tiempo de vida  $\zeta > a$ , es decir

$$\mathcal{E}^{(a)} = \{\omega \in \Omega' : \zeta > a \text{ y } \omega(t) \neq 0 \text{ para toda } 0 < t < \zeta\},$$

y por  $\mathcal{E} = \bigcup_{a>0} \mathcal{E}^{(a)}$  al conjunto de excursiones. Estos conjuntos se consideran con la topología inducida por la topología de Skorokhod en  $\Omega'$ . Con la notación anterior, para  $a > 0$  dada con  $\bar{\Pi}(a) > 0$ , denotemos por  $n(\cdot \mid \zeta > a)$  a la medida de probabilidad sobre  $\mathcal{E}^{(a)}$  asociada a la ley del proceso  $\{X_{g_1(a)+t} : 0 \leq t < \ell_1(a)\}$ . Esta medida de probabilidad se conoce como la *ley de las excursiones de  $X$  con tiempo de vida  $\zeta > a$* .

### 1. Proposición

Sea  $a > 0$  dada con  $\bar{\Pi}(a) > 0$ . Para todo  $b \in (0, a)$  y todo evento medible  $\Lambda \in \mathcal{E}^{(a)}$ ,

$$\bar{\Pi}(a)n(\Lambda \mid \zeta > a) = \bar{\Pi}(b)n(\Lambda \mid \zeta > b).$$

Demostración: De la proposición 2.6 sabemos que

$$\frac{\bar{\Pi}(a)}{\bar{\Pi}(b)} = \mathbb{P}(N_b(g_1(a)) = 0),$$

lo cual nos dice la probabilidad de que la primera excursión con longitud  $\ell > b$  es tal que  $\ell > a$ . Entonces,

$$\frac{\bar{\Pi}(b)}{\bar{\Pi}(a)} n(\cdot, \zeta > a \mid \zeta > b)$$

es la ley de la primera excursión con tiempo de vida  $\zeta > a$  condicional en el evento  $\{N_b(g_1(a)) = 0\}$ . Por la misma proposición 2.6, sabemos que la primera excursión con  $\zeta > a$  es independiente de  $N_b(g_1(a))$ , de modo que la probabilidad condicional anterior es igual a  $n(\cdot \mid \zeta > a)$ . ■

Una consecuencia de la proposición 1 es la existencia de una única medida  $n$  sobre  $\mathcal{E}$ , llamada la *medida de excursiones de  $X$* , la cual satisface

$$n(\Lambda) = \bar{\Pi}(a)n(\Lambda \mid \zeta > a) \quad \text{para } \Lambda \subset \mathcal{E}^{(a)}.$$

En particular obtenemos  $n(\zeta > a) = \bar{\Pi}(a)$ , de modo que la ley de las excursiones de  $X$  con tiempo de vida  $\zeta > a$  se puede pensar como  $n$  condicionada en el evento

$\{\zeta > a\}$ . Otra consecuencia de importancia es saber que la medida de excursiones goza de la propiedad de Markov simple; para precisar lo anterior, dada  $a > 0$ ,

$$g_1(a) + a = \inf\{t \geq a : X_s \neq 0 \text{ para todo } s \in [t - a, t]\}$$

es un tiempo de paro con respecto a  $\{\mathcal{F}_t : t \geq 0\}$ . Por la propiedad fuerte de Markov de  $X$  y la definición de la medida de excursiones, condicional en  $\{\epsilon(a) = x, a < \zeta\}$  (con  $\epsilon$  una excursión genérica y  $\zeta$  su tiempo de vida), bajo  $n$ , el proceso trasladado  $\{\epsilon(t + a) : 0 \leq t \leq \zeta - a\}$  es independiente de  $\{\epsilon(t) : 0 \leq t \leq a\}$  y tiene la misma distribución que  $\{X_t : 0 \leq t < R_0\}$  bajo  $\mathbb{P}_x$ . Dicho de otro modo, el proceso trasladado tiene la misma ley que el proceso de Markov original comenzado en el punto  $x$  y matado en el primer retorno al origen.

Es importante remarcar que la definición de la medida de excursiones depende de la constante  $c$ , pero el cambiar este valor únicamente afecta por un factor multiplicativo.

Previamente se remarcó que los intervalos de excursión son aquellos de la forma  $(L_t^{-1}, L_{t-}^{-1})$  para  $L_t^{-1} < L_{t-}^{-1}$ . Esto sugiere el uso del tiempo local inverso como una escala de tiempo bajo la cual podemos analizar las excursiones de  $X$ . Para ello introducimos el proceso de excursiones de  $X$ , mismo que denotaremos por  $e = \{e_t : t \geq 0\}$ , el cual tomará valores en el espacio de excursiones con un punto aislado  $\Upsilon$ ,  $\mathcal{E} \cup \{\Upsilon\}$ , y es dado por

$$(1) \quad e_t = \begin{cases} \{X_{s+L_t^{-1}} : 0 \leq s < \Delta L_t^{-1}\} & \text{si } L_t^{-1} < L_{t-}^{-1}, \\ \Upsilon & \text{en otro caso.} \end{cases}$$

Para enunciar el siguiente resultado, el cual se debe a Itô (1972), conviene recordar lo que es un proceso Poisson puntual detenido. Dado un proceso Poisson puntual  $\eta = \{\eta_t : t \geq 0\}$ , definimos el tiempo aleatorio  $T_B = \inf\{t \geq 0 : \eta_t \in B\}$  para  $B$  medible. El proceso  $\{\eta_t : 0 \leq t \leq T_B\}$  se conoce como el *proceso Poisson puntual detenido en el primer punto en B*.

## 2. Teorema

- (i) Si 0 es recurrente, entonces  $\{e_t : t \geq 0\}$  es un proceso Poisson puntual con medida característica  $n$ .
- (ii) Si 0 es transitorio entonces  $\{e_t : t \leq L_\infty\}$  es un proceso Poisson puntual con medida característica  $n$ , detenido en el primer punto en  $\mathcal{E}^{(\infty)}$ , el espacio de excursiones con tiempo de vida infinita.

**Demostración:** Comenzamos considerando el caso cuando 0 es recurrente. En vista de la proposición 3.1,  $L_t^{-1}$  es tiempo de paro para cada  $t \geq 0$ , de modo que  $\{\mathcal{H}_t : t \geq 0\}$  es una filtración cuando se define  $\mathcal{H}_t = \mathcal{F}_{L_t^{-1}}$ . Bastará verificar que para cada  $\varepsilon > 0$  y cada conjunto medible  $B \subset \mathcal{E}^{(\varepsilon)}$ , el proceso de conteo  $N_t^B = \#\{0 < s \leq t : e_s \in B\}$ ,  $t \geq 0$ , es un proceso Poisson con respecto a  $\{\mathcal{H}_t : t \geq 0\}$  cuya medida de intensidad es  $n(B)$ . En efecto, de hacerlo, en vista de que si  $B_1, \dots, B_k$  son conjuntos medibles disjuntos entonces los procesos de saltos asociados son independientes debido a que nunca saltan a la vez, deduciremos que la medida aleatoria  $M = \sum_{t \geq 0} \delta_{(t, e_t)}$  es una medida aleatoria de Poisson con intensidad  $\lambda \otimes n$ , donde  $\lambda$  es la medida de Lebesgue en  $[0, \infty)$ .

Observemos que para cada elección de  $s, t \geq 0$ ,  $N_{t+s}^B - N_t^B$  cuenta el número de excursiones de  $X$  en  $B$  que fueron completadas en el intervalo de tiempo  $(L_t^{-1}, L_{t+s}^{-1}]$ . Si ahora introducimos al proceso  $\tilde{X} = X \circ \theta_{L_t^{-1}}$  entonces él es independiente de  $\mathcal{H}_t$  y tiene la misma ley que  $X$  debido a que  $L_t^{-1}$  es un tiempo de paro y en adición es un cero de  $X$  por la parte (iii) de la proposición 3.1. Denotamos por  $\tilde{L}$  y  $\tilde{L}^{-1}$  al tiempo local y tiempo local inverso de  $\tilde{X}$  de manera respectiva. Procediendo como en (3.1), deducimos que  $\tilde{L}_u^{-1} = L_{u+t}^{-1} - L_t^{-1}$  para toda  $u \geq 0$ , y por consiguiente  $N_{t+s}^B - N_t^B = \tilde{N}_s^B$  es el número de excursiones de  $\tilde{X}$  en  $B$  que fueron completadas en el intervalo de tiempo  $(0, \tilde{L}_s^{-1}]$ . Como consecuencia,  $N_{t+s}^B - N_t^B$  tiene la misma distribución que  $N_s^B$  y es independiente de  $\mathcal{H}_t$ . Puesto que los incrementos de  $N^B$  son, casi seguramente, iguales a uno, concluimos que  $N^B$  es un proceso Poisson con respecto a  $\{\mathcal{H}_t : t \geq 0\}$ . Así,  $e$  es un proceso Poisson puntual. A continuación veremos que su medida característica  $\nu$  coincide con  $n$ .

Es sabido, ver por ejemplo el lema 3 de Pardo Millán (s.f.-a) o la proposición 0.2 de Bertoin (1996), que para toda  $u > 0$ , la ley condicional  $\nu(\cdot \mid \mathcal{E}^{(u)})$  es la ley de las excursiones con tiempo de vida  $\zeta > u$ , es decir

$$(2) \quad \frac{\nu(\cdot, \zeta > u)}{\nu(\zeta > u)} = n(\cdot \mid \zeta > u) = \frac{n(\cdot, \zeta > u)}{n(\zeta > u)}.$$

Por otra parte,  $L_{d_1(u)}$  es el instante en el que se observa el primer punto de  $e$  en  $\mathcal{E}^{(u)}$ . Gracias a la proposición 2.6 sabemos que  $L_{d_1(u)}$  tiene distribución exponencial de parámetro  $\bar{\Pi}(u)$ , con lo cual deducimos  $\nu(\zeta > u) = \bar{\Pi}(u)$ . De lo anterior y (2) concluimos que  $\nu$  y  $n$  coinciden en  $\mathcal{E}^{(u)}$ . Puesto que  $\mathcal{E} = \bigcup_{u>0} \mathcal{E}^{(u)}$  obtenemos la igualdad entre  $n$  y  $\nu$ .

Supongamos ahora que 0 es transitorio. Mediante argumentos similares a los ya exhibidos, condicional en  $\{L_\infty > t\}$ , el proceso  $\{e_s : 0 \leq s \leq t\}$  tiene la misma ley que un proceso Poisson puntual restringido al intervalo  $[0, t]$ . A continuación obtendremos la medida característica. Para proseguir, notemos que

$$L_\infty = \inf\{t \geq 0 : N_t^{\mathcal{E}^{(\infty)}} > 0\},$$

de modo que  $N_\infty^{\mathcal{E}^{(\infty)}} = 1$  y  $N_\infty^{\mathcal{E} \setminus \mathcal{E}^{(\infty)}} - N_{L_\infty}^{\mathcal{E} \setminus \mathcal{E}^{(\infty)}} = 0$  casi seguramente. Dicho lo anterior, observemos que dados  $u > 0$ ,  $0 = t_0 < t_1 < \dots < t_n = t$ , además de  $B_1, \dots, B_n \subset \mathcal{E}^{(u)}$  y  $B \subset \mathcal{E}^{(\infty)}$  medibles y ajenos, la propiedad fuerte de Markov nos da la segunda igualdad en

$$\begin{aligned} & \mathbb{P}\left(\{N_{t_i}^{B_j} - N_{t_{i-1}}^{B_j} = k_{ij}, i, j \in \llbracket 1, n \rrbracket\} \cap \{L_\infty > t, N_\infty^B = 1\}\right) \\ &= \mathbb{P}\left(\bigcap_{i=1}^n \left(\{N_{t_i}^{B_j} - N_{t_{i-1}}^{B_j} = k_{ij}, j \in \llbracket 1, n \rrbracket\} \cap \{L_\infty > t_i\}\right) \cap \{L_\infty > t, N_\infty^B = 1\}\right) \\ &= \prod_{i=1}^n \mathbb{P}\left(\{N_{t_i - t_{i-1}}^{B_j} = k_{ij}, j \in \llbracket 1, n \rrbracket\} \cap \{L_\infty > t_i - t_{i-1}\}\right) \mathbb{P}(N_\infty^B = 1). \end{aligned}$$

Debido a que la ley de una excursión con tiempo de vida  $\zeta = \infty$  es dada por  $n(\cdot, \zeta = \infty)/n(\zeta = \infty)$  deducimos inmediatamente

$$\mathbb{P}(N_\infty^B = 1) = \mathbb{P}(e_{L_\infty} \in B) = \frac{n(B)}{n(\zeta = \infty)}.$$

Por otra parte, en  $\{L_\infty > s\}$ ,  $N_s^{B_j} = N_s^{B_j \setminus \mathcal{E}^{(\infty)}}$  casi seguramente, y entonces, al razonar

como en el caso recurrente, deducimos que

$$\begin{aligned} & \mathbb{P}\left(\{N_{t_i}^{B_j} - N_{t_{i-1}}^{B_j} = k_{ij}, i, j \in \llbracket 1, n \rrbracket\} \cap \{L_\infty > t, N_\infty^B = 1\}\right) \\ &= \prod_{i=1}^n \frac{((t_i - t_{i-1})n(B_j \setminus \mathcal{E}^{(\infty)}))^{k_{ij}}}{k_{ij}!} e^{-(t_i - t_{i-1})n(B_j \setminus \mathcal{E}^{(\infty)})} e^{-(t_i - t_{i-1})\bar{\Pi}(\infty)} \frac{n(B)}{n(\zeta = \infty)} \\ &= \prod_{i=1}^n \frac{((t_i - t_{i-1})n(B_j \setminus \mathcal{E}^{(\infty)}))^{k_{ij}}}{k_{ij}!} e^{-(t_i - t_{i-1})n(B_j \setminus \mathcal{E}^{(\infty)})} e^{-t\bar{\Pi}(\infty)} \frac{n(B)}{n(\zeta = \infty)}. \end{aligned}$$

En la primera igualdad se usó que  $L_\infty$  tiene distribución exponencial de parámetro  $\bar{\Pi}(\infty)$ . Lo anterior muestra precisamente lo deseado, es decir que  $\{e_t : 0 \leq t \leq L_\infty\}$  es un proceso Poisson puntual con medida característica  $n$ , detenido en el primer punto en  $\mathcal{E}^{(\infty)}$ . ■

Con este resultado en mente, es posible realizar cálculos usando fórmulas, como la de compensación y la exponencial, en términos de procesos Poisson puntuales. Un par de ejemplos de esto se presentan al final de la sección iv.4 de Bertoin (1996, págs. 119–121).

Antes de concluir con este capítulo, en el cual se introdujeron las nociones de tiempo local y proceso de excursiones, realizaremos algunos comentarios de lo que sucede cuando no suponemos que 0 sea instantáneo o regular.

## 2.A. Puntos de retención y puntos irregulares

A lo largo del capítulo se obtuvieron conexiones entre las excursiones de procesos de Hunt fuera de un punto regular e instantáneo con subordinadores y procesos Poisson puntuales. Cuando el punto en cuestión no es regular o instantáneo, existen conexiones similares, lo cual veremos en esta sección.

Primero discutiremos el caso en el cual 0 es un punto de retención. Consideremos la sucesión de salidas y retornos al 0:  $R_0 < S_1 < R_1 < \dots$ , donde  $R_0 = 0$  y

$$R_n = \inf\{t > S_n : X_t = 0\} \quad \text{y} \quad S_{n+1} = \inf\{t > R_n : X_t \neq 0\}.$$

Por la propiedad fuerte de Markov aunada a que  $X$  tiene trayectorias càd,  $X_{S_1} \neq 0$ ,  $S_1$  tiene distribución exponencial y es independiente de la primera excursión fuera de 0,  $\{X_{S_1+t} : 0 \leq t < R_1 - S_1\}$ . Ahora, en el evento  $\{R_1 < \infty\}$  se tiene  $X_{R_1} = 0$ , y podemos repetir el argumento gracias a la propiedad fuerte de Markov. Consecuentemente, podemos expresar al conjunto de ceros de  $X$  como

$$\mathcal{Z} = \bigcup_{n \geq 0} [R_n, S_{n+1}).$$

Luego, existe un funcional aditivo, en el sentido dado por (2.5), que es continuo y crece precisamente en  $\text{cl}(\mathcal{Z})$ , a saber

$$A_t = \int_0^t \mathbb{1}_{\{X_s=0\}} ds, \quad t \geq 0.$$

Así, podemos definir el tiempo local del 0 como cualquier proceso  $L = \{L_t : t \geq 0\}$  que satisfaga  $\partial L = A$  para alguna constante  $\partial > 0$ . Es sabido que la inversa continua por la derecha  $L^{-1}$  de  $L$  es continua salvo en los intervalos de constancia de  $L$ , es decir que es continua salvo en el conjunto  $\{L_{S_i} : i \geq 1\}$ . Dado lo anterior, es fácil

convencerse de que se satisface un resultado análogo a la proposición 3.1 y que  $L^{-1}$  es un subordinador, matado a cierta tasa en el caso transitorio, con deriva  $\partial$ . Por lo mismo, el proceso de excursiones (4.1) es un proceso Poisson puntual, parado en la primera vez que  $e_t \in \mathcal{E}^{(\infty)}$ . Finalmente, notando que en este caso  $\bar{\Pi}(a) \rightarrow \infty$  conforme  $a \rightarrow 0$  (revisar el argumento con el cual se demostró el lema 2.4), la medida característica del proceso puntual es finita, y entonces es proporcional a la ley de la primera excursión de  $X$  fuera de 0,  $\{X_{S_1+t} : 0 \leq t < R_1 - S_1\}$ .

Consideremos ahora que 0 es irregular. Tomemos ahora la sucesión de retornos al cero  $\{R_n : n \in \mathbb{N}\}$  definida formalmente por  $R_0 = 0$  y  $R_{n+1} = \inf\{t > R_n : X_t = 0\}$ . En este caso tenemos  $R_1 \in (0, \infty]$  casi seguramente. Condicional en  $\{R_1 < \infty\}$ , la propiedad fuerte de Markov nos dice que el proceso  $X \circ \theta_{R_1}$  es independiente de la primera excursión  $\{X_t : 0 \leq t \leq R_1\}$  y tiene la misma ley que  $X$ . Repitiendo el argumento observamos que  $\{R_n : n \in \mathbb{N}\}$  es una caminata aleatoria creciente, matada en un tiempo aleatorio geométrico en el caso transitorio. Lo más natural sería definir al tiempo local al tiempo  $t$  como el número de retornos al origen antes de  $t$ . No obstante, con esta definición obtenemos un proceso con valores enteros cuyo inverso continuo por la derecha coincide con  $\{R_n : n \in \mathbb{N}\}$ , es decir que es un proceso en tiempo discreto. Puesto que sería más conveniente tener una relación entre subordinadores y excursiones, como antes, definiremos al tiempo local mediante un artificio. Recordando que al subordinar una caminata aleatoria por un proceso Poisson obtenemos un proceso Poisson compuesto, una instancia particular de un proceso de Lévy, consideraremos una sucesión  $\{\tau_n : n \in \mathbb{N}_0\}$  de variables aleatorias independientes y con distribución exponencial de mismo parámetro para todas. Esta sucesión además será independiente de  $X$ . Llamamos tiempo local de  $X$  en 0 al proceso  $L = \{L_t : t \geq 0\}$  dado por

$$L_t = \sum_{i \geq 0} \tau_i \mathbb{1}_{\{R_i < t\}} = \sum_{i=0}^{n(t)} \tau_i, \quad \text{donde } n(t) := \max\{i : R_i < t\}.$$

Por construcción,  $L$  crece precisamente en  $\mathcal{Z}$ . Sin embargo, este proceso no es adaptado a la filtración  $\{\mathcal{F}_t : t \geq 0\}$  y es únicamente continuo por la derecha. La discontinuidad únicamente causa intervalos de constancia en la inversa, lo cual no es problemático. El problema de medibilidad lo podemos evitar reemplazando la filtración  $\{\mathcal{F}_t : t \geq 0\}$  por  $\{\mathcal{F}'_t : t \geq 0\}$ , donde

$$\mathcal{F}'_t = \mathcal{F}_t \vee \sigma(L_s, 0 \leq s \leq t).$$

Con ello es posible deducir que la inversa continua por la izquierda es un subordinador, matado en un tiempo exponencial en el caso transitorio. Más aún, el proceso de excursiones vuelve a ser un proceso Poisson puntual, detenido en el caso transitorio, con medida característica proporcional a la ley de la primera excursión de  $X$  fuera de 0.

# Procesos de Lévy de variación finita

# 3

## 3.1. Supuestos y preliminares

A lo largo del capítulo supondremos que  $X$  es un proceso de Lévy sobre  $\mathbb{R}$  *espectralmente positivo* y de *variación finita*, es decir que su medida de Lévy  $\Pi$  tiene soporte en  $(0, \infty)$  y es tal que

$$\int_{(0, \infty)} (r \wedge 1) d\Pi(r).$$

Además supondremos que el proceso de Lévy comienza en 0, i.e.  $\mathbb{P} = \mathbb{P}_0$ .

Por la descomposición de Lévy–Itô, el proceso se puede escribir como la suma de un subordinador  $Y_t = \sum_{0 \leq s \leq t} \Delta X_s$  y una parte de deriva  $Z_t = \beta t$  con  $\beta \in \mathbb{R}$ . Ahora bien, como

$$\sum_{s: \Delta X_s > 0} \delta_{(s, \Delta X_s)}$$

es un proceso Poisson puntual con intensidad  $ds d\Pi(r)$ , vemos que

$$\mathbb{E}\left[\sum_{0 \leq s \leq t} \Delta X_s\right] = t \int_0^\infty r d\Pi(r)$$

y, más aún, para  $\lambda > 0$ ,

$$\mathbb{E}[\exp\{-\lambda X_t\}] = \exp\{t\Psi(\lambda)\},$$

donde

$$\Psi(\lambda) = -\beta\lambda + \int_0^\infty (e^{-\lambda r} - 1) d\Pi(r).$$

Nosotros estaremos interesados en el caso donde  $X$  es recurrente o deriva a  $-\infty$ , situación que ocurre si, solamente si,  $\mathbb{E}[X_t] \leq 0$ , condición que es equivalente a

$$\beta + \int_0^\infty r d\Pi(r) \leq 0.$$

Esto en particular implica  $\int_0^\infty r d\Pi(r) < \infty$ . Ponemos  $\alpha = -\beta \geq 0$  y de hecho supondremos que  $\Pi \neq 0$ , de modo que  $\alpha > 0$ .

En lo que resta, serán necesarios algunos resultados del comportamiento de  $X$ . Lo primero que afirmamos es que 0 es irregular para  $(0, \infty)$ . En efecto, si definimos el tiempo de paro  $\gamma_{(t)} = \inf\{s > t : X_s > X_t\}$  entonces por los incrementos independientes y estacionarios de  $X$ ,  $\mathbb{P}(\gamma_{(t)} > t)$  es independiente de  $t$ . En adición,  $\gamma_{(t)} > t$  si y sólo si la derivada de  $s \mapsto X_s$  existe en  $s = t$  y es negativa. Puesto que la derivada es igual a  $-\alpha$  para casi todo punto con respecto a  $dt$ , obtenemos la irregularidad. Lo anterior en particular nos dice que el supremo del proceso  $X$  únicamente incrementa mediante saltos.

Denotamos por  $\gamma = \gamma_{(0,\infty)}$  el tiempo de entrada a  $(0, \infty)$ . Entonces la distribución de  $(X_\gamma, \Delta X_\gamma)$  está dada por

$$(1) \quad \mathbb{E}[\mathbb{1}_{\{\gamma < \infty\}} f(X_\gamma, \Delta X_\gamma)] = \frac{1}{\alpha} \int_0^\infty \int_0^y f(x, y) dx d\Pi(y).$$

Para demostrar (1) primero hacemos notar que  $\Delta X_\gamma > 0$  por lo dicho en el párrafo anterior. Ahora bien, observemos que para  $B \subset (-\infty, 0)$  boreliano y  $\varepsilon > 0$ , si definimos  $\gamma_{(-\varepsilon)} = \inf\{t > 0 : X_t = -\varepsilon\}$ , al notar que  $X_{\gamma_{(-\varepsilon)}} = -\varepsilon$ , tendremos

$$\begin{aligned} \mathbb{E}\left[\int_0^\gamma \mathbb{1}_{B-\varepsilon}(X_s) ds\right] &= \mathbb{E}\left[\mathbb{1}_{\{\gamma_{(-\varepsilon)} < \gamma\}} \int_{\gamma_{(-\varepsilon)}}^\gamma \mathbb{1}_{B-\varepsilon}(X_s) ds\right] \\ &= \mathbb{E}\left[\mathbb{1}_{\{\gamma_{(-\varepsilon)} < \gamma\}} \int_{\gamma_{(-\varepsilon)}}^\gamma \mathbb{1}_B(X_{s-\gamma_{(-\varepsilon)}+\gamma_{(-\varepsilon)}} - X_{\gamma_{(-\varepsilon)}}) ds\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\{0 < \gamma - \gamma_{(-\varepsilon)}\}} \int_0^{\gamma - \gamma_{(-\varepsilon)}} \mathbb{1}_B(X_{s+\gamma_{(-\varepsilon)}} - X_{\gamma_{(-\varepsilon)}}) ds \mid \mathcal{F}_{\gamma_{(-\varepsilon)}}\right]\right] \\ &= \mathbb{E}_{-\varepsilon}\left[\int_0^\gamma \mathbb{1}_B(X_s + \varepsilon) ds\right] \\ &= \mathbb{E}\left[\int_0^\gamma \mathbb{1}_B(X_s) ds\right], \end{aligned}$$

es decir que la medida  $B \mapsto \mathbb{E}[\int_0^\gamma \mathbb{1}_B(X_s) ds]$  es invariante bajo traslaciones. Por ende, existe una constante  $c > 0$  tal que

$$(2) \quad \mathbb{E}\left[\int_0^\gamma \mathbb{1}_B(X_s) ds\right] = c\lambda(B),$$

donde  $\lambda$  es la medida de Lebesgue. Luego, notando que

$$\mathbb{1}_{\{\gamma < \infty\}} = \sum_{s: \Delta X_s > 0} \mathbb{1}_{\{s \leq \gamma\}} \mathbb{1}_{\{\Delta X_s > -X_{s-}\}}$$

podemos usar la fórmula de compensación para obtener la tercera igualdad en

$$\begin{aligned} \mathbb{E}[\mathbb{1}_{\{\gamma < \infty\}} f(X_\gamma, \Delta X_\gamma)] &= \mathbb{E}\left[\sum_{s: \Delta X_s > 0} \mathbb{1}_{\{s \leq \gamma\}} \mathbb{1}_{\{\Delta X_s > -X_{s-}\}} f(X_s, \Delta X_s)\right] \\ &= \mathbb{E}\left[\sum_{s: \Delta X_s > 0} \mathbb{1}_{\{s \leq \gamma\}} \mathbb{1}_{\{\Delta X_s > -X_{s-}\}} f(X_{s-} + \Delta X_s, \Delta X_s)\right] \\ (3) \quad &= \mathbb{E}\left[\int_0^\infty \int_{\mathbb{R}} \mathbb{1}_{\{s \leq \gamma\}} \mathbb{1}_{\{r > -X_{s-}\}} f(X_{s-} + r, r) d\Pi(r) ds\right] \\ &= \mathbb{E}\left[\int_0^\gamma \int_{(-X_{s-}, \infty)} f(X_{s-} + r, r) d\Pi(r) ds\right] \\ &= c \int_{-\infty}^0 \int_{(-x, \infty)} f(x + r, r) d\Pi(r) dx. \end{aligned}$$

En vista de lo desarrollado, obtendremos (1) si verificamos que  $c = \alpha^{-1}$ . Al considerar  $f = 1$  vemos que  $\mathbb{P}(\gamma < \infty) = c \int_0^\infty r d\Pi(r)$ . En consecuencia, como  $\mathbb{P}(\gamma = 1) = 1$  y  $\alpha = \int_0^\infty r d\Pi(r)$  en el caso recurrente, obtenemos lo deseado para este caso. En el

caso transitorio notamos primero que

$$\begin{aligned}
 \mathbb{E}\left[\int_0^\infty \mathbb{1}_B(X_s) ds\right] &= \mathbb{E}\left[\mathbb{1}_{\{\gamma < \infty\}} \int_0^\gamma \mathbb{1}_B(X_s) ds\right] + \mathbb{E}\left[\mathbb{1}_{\{\gamma < \infty\}} \int_\gamma^\infty \mathbb{1}_B(X_s) ds\right] \\
 &\quad + \mathbb{E}\left[\mathbb{1}_{\{\gamma = \infty\}} \int_0^\infty \mathbb{1}_B(X_s) ds\right] \\
 &= \mathbb{E}\left[\int_0^\gamma \mathbb{1}_B(X_s) ds\right] + \mathbb{E}\left[\mathbb{1}_{\{\gamma < \infty\}} \int_\gamma^\infty \mathbb{1}_B(X_s) ds\right] \\
 &= \mathbb{E}\left[\int_0^\gamma \mathbb{1}_B(X_s) ds\right] + \mathbb{P}(\gamma < \infty) \mathbb{E}\left[\int_0^\infty \mathbb{1}_B(X_s) ds\right] \\
 &= \frac{c}{\mathbb{P}(\gamma = \infty)} \lambda(B),
 \end{aligned}$$

donde hemos usado la propiedad fuerte de Markov en la tercera igualdad. Por otra parte, también es sencillo verificar directamente, probando la invarianza bajo traslaciones, que la medida  $B \mapsto \mathbb{E}[\int_0^\infty \mathbb{1}_B(X_s) ds]$  es proporcional a la medida de Lebesgue. Debido a que  $\mathbb{E}[X_1] = -\alpha + \int_0^\infty r d\Pi(r)$ ,

$$dv(x) = \int_0^\infty \mathbb{P}(X_t \in dx) e^{-t} dt$$

define una medida de probabilidad con  $\int_{\mathbb{R}} x dv(x) = -\alpha + \int_0^\infty r d\Pi(r)$  y transformada de Fourier:

$$\mathcal{F}v(s) = \int_{\mathbb{R}} e^{is} dv(x) = \int_0^\infty e^{-t} e^{t\Psi(-is)} dt = \frac{1}{1 - \Psi(-is)}.$$

Entonces, si para  $x \in \mathbb{R}$  y  $B \in \mathcal{B}(\mathbb{R})$  definimos  $V(x, B) = \mathbb{E}_x[\sum_{n=0}^\infty \mathbb{1}_B(W_n)]$ , con  $W = \{W_n : n \geq 0\}$  una caminata aleatoria con distribución de saltos  $v$ , entonces

$$\mathcal{F}V(0, \cdot)(s) = \sum_{n=0}^\infty \frac{1}{(1 - \Psi(-is))^n} = 1 + \frac{1}{-\Psi(-is)}.$$

Al invertir la transformada de Fourier obtenemos

$$V(0, B) = \delta_0 + \mathbb{E}\left[\int_0^\infty \mathbb{1}_B(X_s) ds\right],$$

y por consiguiente

$$V(x, B) = \delta_x + \mathbb{E}_x\left[\int_0^\infty \mathbb{1}_B(X_s) ds\right].$$

En particular, si  $x > 0$  y  $B \subset (-\infty, 0)$ ,

$$V(x, B) = \mathbb{E}\left[\int_0^\infty \mathbb{1}_B(X_s) ds\right].$$

Entonces, por el teorema 3.1 y la proposición 3.4 de la sección 5.3 de Revuz (1984) obtenemos

$$\mathbb{E}\left[\int_0^\infty \mathbb{1}_B(X_s) ds\right] = \lim_{x \rightarrow \infty} V(x, B) = \frac{1}{\alpha - \int_0^\infty r d\Pi(r)} \lambda(B).$$

De lo anterior, como  $\mathbb{P}(\gamma = \infty) = 1 - c \int_0^\infty r d\Pi(r)$ , obtenemos

$$\frac{1}{\alpha - \int_0^\infty r d\Pi(r)} = \frac{c}{1 - c \int_0^\infty r d\Pi(r)} = \frac{1}{1/c - \int_0^\infty r d\Pi(r)},$$

de donde concluimos  $c = 1/\alpha$ , como era requerido.

Denotamos por  $I_t = \inf_{0 \leq s \leq t} X_s$  y  $S_t = \sup_{0 \leq s \leq t} X_s$  el ínfimo y supremo del proceso  $X$  hasta el tiempo  $t$ . De acuerdo a la proposición VI.1 de Bertoin (1996), los procesos reflejados  $X - I$  y  $S - X$  son procesos de Markov. Por los supuestos realizados sobre  $X$  se sigue que 0 es un punto regular y recurrente para  $X - I$ . Ahora, del hecho de que 0 es irregular para  $(0, \infty)$  obtenemos  $\mathbb{P}(X_t = I_t) > 0$  para toda  $t \geq 0$  y en consecuencia, como  $S_t$  y  $X_t - I_t$  tienen la misma distribución para  $t \geq 0$  fija,  $\mathbb{P}(S_t = 0) > 0$  para toda  $t \geq 0$ . Por las observaciones de la sección 2.A, un tiempo local de 0 para  $X - I$  es

$$A_t = \int_0^t \mathbb{1}_{\{X_s = I_s\}} ds.$$

A la medida de excursiones de  $X - I$  la denotamos por  $N$ . Para obtener su forma explícita hacemos notar primero que las excursiones de  $X - I$  deben comenzar con un salto, debido a que  $X$  únicamente crece mediante saltos. Por lo anterior,

$$\begin{aligned} \int_{\mathcal{E}} f(\epsilon_0) dN(\epsilon) &= \mathbb{E} \left[ \int_0^{\tau_{(1)}} \int_{\mathcal{E}} f(\epsilon_0) dN(\epsilon) dA_s \right] \\ &= \mathbb{E} \left[ \sum_s \mathbb{1}_{\{A_s^{-1} \leq 1\}} f(\Delta X_{A_s^{-1}}) \right] = \mathbb{E} \left[ \sum_{\nu \leq \tau_{(1)}} f(\Delta X_\nu) \right] \\ &= \mathbb{E}[\tau_{(1)}] \int_0^\infty f(r) d\Pi(r), \end{aligned}$$

donde  $\tau_{(1)} = \inf\{t > 0 : A_t = 1\}$ . Luego, si denotamos por  $\mathbb{P}_r^0$  a la ley del proceso  $X$  comenzado en  $r > 0$  y parado al tiempo  $T_0 = \inf\{s \geq 0 : X_s = 0\}$ , deducimos que

$$N = \mathbb{E}[\tau_{(1)}] \int_0^\infty \mathbb{P}_r^0 d\Pi(r).$$

Con los cálculos realizados obtenemos algo distinto a Le Gall y Le Jan (1998, pág. 222), donde afirman que  $N$  es tal que

$$(4) \quad N(f(\epsilon_0)) = \int f(r) d\Pi(r).$$

Para obtener esta relación es necesaria una normalización del tiempo local  $L$  para la cual

$$\mathbb{E}[\tau_{(1)}^L] = 1.$$

A la medida de excursiones asociada a tal tiempo local la denotaremos por  $N$  y por construcción satisfará (4).

## 3.2. Representación como sistema de colas

El proceso de Lévy  $X$  con el cual trabajaremos a lo largo de este capítulo puede entenderse como la evolución de una cola. Esta interpretación la describiremos a lo largo de esta sección; ella nos dará una estructura de ramificaciones que estudiaremos en apartados posteriores.

Primero consideraremos el caso en el cual  $\Pi$  es una medida finita, de modo que si definimos  $Y_t = \sum_{s \leq t} \Delta X_s$ ,  $Y = \{Y_t : t \geq 0\}$  es un proceso de Poisson compuesto. Los saltos de este tienen distribución  $\Pi/\Pi(\mathbb{R}_+)$  y sus tiempos entre saltos tienen

distribución exponencial de parámetro  $\Pi(\mathbb{R}_+)$ . En consecuencia  $X$  es elemento del conjunto  $\mathcal{T} \subset \mathbb{D}(\mathbb{R}_+, \mathbb{R})$  de funciones  $\omega = \{\omega(t) : t \geq 0\}$  de la forma  $\omega(t) = u(t) - \alpha t$ , donde  $u$  es una suma de saltos discretos.

Dicho lo anterior, interpretaremos la trayectoria  $t \mapsto X_t$  como la evolución en el tiempo de una cola tipo LIFO<sup>1</sup> con un único servidor, cuya tasa de servicio es  $\alpha$ . Un salto de  $X$  al tiempo  $t$  indica que llegó un nuevo cliente que requiere un servicio  $\Delta X_t$ . Por el tipo de cola, en cuanto llega el nuevo cliente, el servidor procede a atenderle y su servicio quedará completo al tiempo  $t + \alpha^{-1}\Delta X_t$  salvo que otro cliente llegue antes de que el servicio termine. En cuanto termine el servicio de un cliente, el servidor procede a atender al último cliente en llegar cuyo servicio no haya terminado. De esta descripción se deduce fácilmente que un cliente que llega al tiempo  $s > 0$  estará en el sistema al tiempo  $t > s$  si, únicamente si,

$$X_{s-} < \inf_{s \leq r \leq t} X_r,$$

en cuyo caso su servicio remanente es dado por

$$\inf_{s \leq r \leq t} X_r - X_{s-}.$$

En vista de esta descripción, si  $N_t$  es el número de clientes que se encuentran en la cola al tiempo  $t$  entonces

$$N_t = \#\{s \in [0, t] : X_{s-} < \inf_{s \leq r \leq t} X_r\}.$$

Denotamos por  $\mathcal{J}_t := \{s_t^1 \leq \dots \leq s_t^{N_t}\}$  al conjunto de tiempos de llegada de tales clientes. Para los mismos clientes, denotamos por  $\rho_t(s_t^1), \dots, \rho_t(s_t^{N_t})$  sus servicios remanentes al tiempo  $t$ ; es decir  $\rho_t(s) = \inf_{s \leq r \leq t} X_r - X_{s-}$ . Observemos que el servidor se encuentra inactivo al tiempo  $t$  si y sólo si  $X_t = I_t$ . En caso de que el servidor se encuentre activo, cuando  $X_t > I_t$  su carga se encuentra dada precisamente por  $X_t - I_t$ , lo cual nos da la relación

$$\sum_{s \in \mathcal{J}_t} \rho_t(s) = X_t - I_t.$$

Una representación gráfica de esta última igualdad se muestra en la figura 1. Por la misma razón de que el servidor se encuentre inactivo precisamente cuando  $X = I$ ,

$$-\frac{1}{\alpha}I_t = \int_0^t \mathbb{1}_{\{X_s = I_s\}} ds$$

representa el tiempo que el servidor se encuentra inactivo hasta el tiempo  $t$ .

En el caso general, cuando  $\Pi$  puede ser una medida infinita,  $X$  pertenece al conjunto  $\overline{\mathcal{T}} \subset \mathbb{D}(\mathbb{R}_+, \mathbb{R})$  de trayectorias  $\omega$  de la forma  $\omega(t) = u(t) - \alpha t$ , donde ahora  $u(t)$  es una suma de saltos. La interpretación antes dada sigue siendo válida. Seguimos denotando por  $\mathcal{J}_t$  al conjunto

$$\{s \in [0, t] : \rho_t(s) = \inf_{s \leq r \leq t} X_r - X_{s-} > 0\}$$

y por  $N_t$  a su cardinalidad, que ahora puede ser infinita.

<sup>1</sup>*Last in, first out*, también conocida como LCFS por *Last come, first served*.

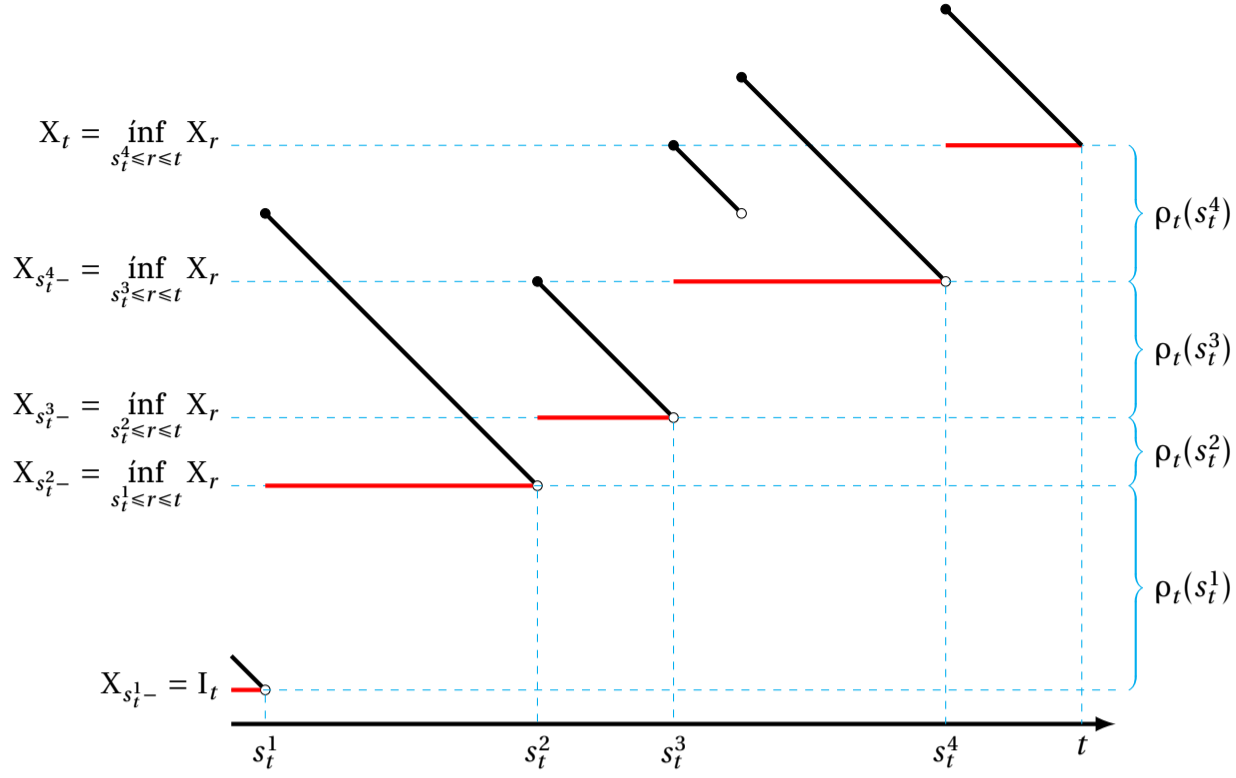


Figura 1: Descomposición de  $X_t - I_t$  bajo la interpretación de colas tipo LIFO. En rojo se muestra la trayectoria  $s \in [0, t] \mapsto \inf_{s \leq r \leq t} X_r$ .

Para estudiar el caso  $\Pi(\mathbb{R}_+) = \infty$  usaremos un argumento de truncamiento. Para cada  $\varepsilon > 0$  definimos  $X^\varepsilon = \{X_t^\varepsilon : t \geq 0\}$  mediante

$$X_t^\varepsilon = \sum_{0 \leq s \leq t} \Delta X_s 1_{\{\Delta X_s > \varepsilon\}} - \alpha t.$$

De acuerdo a la interpretación dada, la trayectoria  $t \mapsto X_t^\varepsilon$  corresponde entonces a la evolución de una cola tipo LIFO donde únicamente se aceptan clientes cuyo servicio sea mayor que  $\varepsilon$ . Hacer  $\varepsilon$  menor significa agregar clientes con servicios menores. Entonces  $X^\varepsilon$  pertenece a  $\mathcal{T}$  y en adición converge a  $X$ , conforme  $\varepsilon \rightarrow 0$ , de manera uniforme sobre compactos. Más aún,  $\mathcal{J}_t^\varepsilon$  crece a  $\mathcal{J}_t$ , implicando que  $N_t^\varepsilon$  crece a  $N_t$  conforme  $\varepsilon \rightarrow 0$ . Asimismo tenemos que  $\rho_t^\varepsilon(s)$  crece a  $\rho_t(s)$  y en consecuencia

$$X_t - I_t = \lim_{\varepsilon \rightarrow 0} (X_t^\varepsilon - I_t^\varepsilon)$$

representa la carga del servidor al tiempo  $t$ . Por lo anterior, podemos pensar en que  $t \mapsto (\mathcal{J}_t, \rho_t)$  describe la evolución de una cola generalizada donde el número total de clientes en el sistema puede ser infinita.

### 3.3. Estructura de ramificación

A lo largo de esta sección consideraremos las osbervaciones hechas en la sección aplicadas a cualquier elemento de  $\mathcal{T}$  o  $\overline{\mathcal{T}}$ . Para obtener información adicional usaremos la estructura probabilista de  $X$ , comenzando con el siguiente resultado.

#### 1. Lema

Para cada  $t > 0$ ,  $\mathbb{P}(N_t < \infty) = 1$ . Más aún, con probabilidad uno, para cada  $t > 0$  los elementos de  $\mathcal{J}_t$  conforman una sucesión estrictamente creciente  $s_t^1 < s_t^2 < \dots$ . Esta

*sucesión es finita si  $t$  es un tiempo de salto para  $X$ .*

Demostración: Para cualesquiera  $0 \leq s \leq t$  consideramos  $\widehat{X}_s^{(t)} = X_t - X_{(t-s)-}$  y  $\widehat{S}_s^{(t)} = \sup_{0 \leq r \leq s} \widehat{X}_r^{(t)}$ . Debido a que tenemos las igualdades

$$\widehat{S}_u^{(t)} = X_t - \inf_{t-u \leq r \leq t} X_{r-} \quad \text{y} \quad \widehat{S}_{u-}^{(t)} = X_t - \inf_{t-u < r \leq t} X_{r-},$$

$t - u \in \mathcal{J}_t$  si, sólo si,  $\widehat{S}_u^{(t)} > \widehat{S}_{u-}^{(t)}$ .

Ahora, como el conjunto  $\{s > 0 : S_s > S_{s-}\}$  es discreto,  $\{s \in (0, t] : S_s > S_{s-}\}$  es finito casi seguramente para cada  $t$ . Como  $\{\widehat{S}_s^{(t)} : 0 \leq s \leq t\}$  tiene la misma distribución que  $\{S_s : 0 \leq s \leq t\}$ , la observación realizada en el párrafo anterior implica  $N_t < \infty$  casi seguramente. Luego,  $N_r$  es finito para todo racional positivo  $r$  con probabilidad uno.

Para verificar la segunda afirmación del enunciado, observemos que para  $s, s' \in \mathcal{J}_t$ ,  $s < s'$  precisamente cuando  $X_{s-} < X_{s'-}$ . Así, deduciremos el resultado al probar que, con probabilidad uno, para cualesquiera  $t > 0$  y  $\varepsilon > 0$ ,  $\{s \in \mathcal{J}_t : X_{s-} < X_t - \varepsilon\}$  es finito. Con este cometido en mente, usando la continuidad por la derecha de las trayectorias de  $X$  notamos que este conjunto está contenido en  $\mathcal{J}_r$  para algún racional  $r > t$ . En caso que  $t$  sea un tiempo de salto, entonces existe  $\varepsilon > 0$  para el cual  $\mathcal{J}_t = \{s \in \mathcal{J}_t : X_{s-} < X_t - \varepsilon\}$ , lo cual completa la demostración. ■

Para cada  $p \geq 0$  definimos el proceso continuo  $Z^p = \{Z_t^p : t \geq 0\}$  mediante

$$Z_t^p = \int_0^t \mathbb{1}_{\{N_s = p\}} ds.$$

Gracias al lema 1 obtenemos  $\sum_{p \geq 0} Z_t^p = t$ . En adición consideramos  $Z^{\varepsilon, p}$  asociada a proceso de Lévy truncado  $X^\varepsilon$ . Por el teorema de convergencia dominada,  $Z_t^{\varepsilon, p} \rightarrow Z_t^p$  conforme  $\varepsilon \rightarrow 0$ . Lo anterior, aplicado a  $p = 0$  nos dice

$$Z_t^0 = \lim_{\varepsilon \rightarrow 0} Z_t^{\varepsilon, 0} = \lim_{\varepsilon \rightarrow 0} \left( -\frac{1}{\alpha} I_t^\varepsilon \right) = -\frac{1}{\alpha} I_t,$$

pues  $X_t = I_t$  sucede si, únicamente si,  $N_t = 0$ .

Para cada  $x > 0$  consideramos  $\tau_x = \{t \geq 0 : X_t = -x\}$  y notamos que, con probabilidad uno,  $\tau_x = \{t \geq 0 : X_{t-} = -x\}$ , de modo que, casi seguramente,  $\tau_x^\varepsilon \uparrow \tau_x$  conforme  $\varepsilon \downarrow 0$ . Entonces  $Z_{\tau_x}^0 = x/\alpha$ . A continuación veremos cuál es la ley de proceso  $\{Z_{\tau_x}^p : p \geq 0\}$  dada  $x > 0$ .

## 2. Proposición

*El proceso  $\{Z_{\tau_x}^p : p \geq 0\}$  es una cadena de Markov en  $\mathbb{R}_+$  cuyo kernel de transición  $P(u, dv)$  queda determinado por su transformada de Laplace,*

$$\int_{\mathbb{R}_+} e^{-\lambda v} P(u, dv) = \exp \left\{ -u \int_0^\infty \left( 1 - \exp \left\{ -\frac{\lambda}{\alpha} r \right\} \right) d\Pi(r) \right\}.$$

## 3. Observación

Observemos que la proposición 2 nos dice que el kernel  $P$  satisface la propiedad

de ramificación, es decir

$$P(u, \cdot) * P(u', \cdot) = P(u + u', \cdot).$$

Demostración de la proposición 2: Supondremos primero que  $\Pi(\mathbb{R}_+) < \infty$  y obtendremos el resultado general al considerar el límite conforme  $\varepsilon \rightarrow 0$ . Bajo el supuesto de la medida finita sean  $U_0 = 0$  y  $T_1 < U_1 < T_2 < U_2 < \dots$  la sucesión de tiempos de paro definidos mediante

$$T_n = \inf\{t > U_{n-1} : X_{t-} = I_t, \Delta X_t > 0\} \quad \text{y} \quad U_n = \inf\{t > T_n : X_t = X_{T_n-}\}$$

para  $n \in \mathbb{N}$ . Notemos que, por construcción,

$$\{T_1, T_2, \dots\} = \{t > 0 : X_{t-} = I_t, \Delta X_t > 0\},$$

mientras que la propiedad fuerte de Markov implica que la colección de variables aleatorias  $\{T_n - U_{n-1} : n \geq 1\}$  es independiente e idénticamente distribuida con distribución común exponencial de parámetro  $\Pi(\mathbb{R}_+)$ . Por ende, la variable aleatoria

$$j_x := \sup\left\{n : \sum_{i=1}^n (T_i - U_{i-1}) \leq \frac{x}{\alpha}\right\}$$

tiene distribución Poisson de parámetro  $x\Pi(\mathbb{R}_+)/\alpha$ . De la misma construcción obtenemos  $T_n \leq \tau_x$  si y sólo si  $n \leq j_x$ .

En adición, por la propiedad fuerte de Markov, los procesos

$$X^i = \{X_{T_i+t} - X_{T_i} : 0 \leq t \leq U_i - T_i\}, \quad i \in \mathbb{N},$$

son independientes y tienen la misma distribución  $\{X_t : t \leq \tau_\xi\}$ , donde  $\xi$  es una variable aleatoria independiente de  $X$  con distribución  $\Pi(\cdot)/\Pi(\mathbb{R}_+)$ . Por la misma propiedad fuerte de Markov, estos procesos también son independientes de las variables aleatorias  $\{T_n - U_{n-1} : n \geq 1\}$  y en consecuencia de  $j_x$ . Denotamos por  $N^i$  al análogo de  $X$  para  $X^i$ , es decir el proceso que cuenta el número de clientes en el sistema usando la interpretación de  $X^i$  como la evolución de una cola tipo LIFO. De la construcción realizada tendremos  $N_t = 1 + N_{t-T_i}^i$  si  $t \in [T_i, U_i)$ , y  $N_t = 0$  si  $t \notin \bigcup_i [T_i, U_i)$ , de donde

$$\{Z_{\tau_x}^p : p \in \mathbb{N}\} \stackrel{d}{=} \left\{ \sum_{i=1}^{j_x} Z^{i,p-1} : p \in \mathbb{N} \right\},$$

donde  $Z^i = \{Z^{i,k} : k \in \mathbb{N}_0\}$  son procesos independientes entre sí, independientes de  $j_x$  y que tienen la misma distribución que  $Z_{\tau_\xi}$ . Usando la aditividad de las leyes de  $Z_{\tau_x}$ , i.e.  $Z_{\tau_x} + Z_{\tau_y} \stackrel{d}{=} Z_{\tau_{x+y}}$ , vemos que

$$\left\{ \sum_{i=1}^{j_x} Z^{i,p-1} : p \in \mathbb{N} \right\} \stackrel{d}{=} \{Z_{\tau_U}^{p-1} : p \in \mathbb{N}\},$$

donde  $U$  es independiente de  $X$  y tiene la misma distribución que la suma de  $j_x$  copias independientes de  $\xi$ , es decir

$$\begin{aligned} \mathbb{E}[e^{-\lambda U}] &= \mathbb{E}\left[\left(\frac{1}{\Pi(\mathbb{R}_+)} \int_0^\infty e^{-\lambda r} d\Pi(r)\right)^{j_x}\right] \\ &= \exp\left\{\frac{x}{\alpha} \Pi(\mathbb{R}_+) \left(\frac{1}{\Pi(\mathbb{R}_+)} \int_0^\infty e^{-\lambda r} d\Pi(r) - 1\right)\right\} \\ &= \exp\left\{\frac{x}{\alpha} \int_0^\infty (e^{-\lambda r} - 1) d\Pi(r)\right\} = \exp\left\{-\frac{x}{\alpha} \int_0^\infty (1 - e^{-\lambda r}) d\Pi(r)\right\}. \end{aligned}$$

Concluimos al hacer notar que  $Z_{\tau_U}^0 = U/\alpha$  y  $Z_{\tau_x}^0 = x/\alpha$ . ■

Antes de estudiar el proceso  $\rho$ , haremos un par de observaciones. Para la primera supondremos que  $\Pi(\mathbb{R}_+) < \infty$ . Denotamos por  $J_t^p$  el número de saltos del proceso  $N$  de  $p$  a  $p+1$  antes del tiempo  $t$ , de modo que  $J_{\tau_x}^0 = j_x$  con la notación de la demostración anterior. Dicho de otro modo,  $J_t^p$  representa el número de clientes que llegan antes del tiempo  $t$ , en tiempos donde la cantidad de clientes en el sistema es  $p$ . Afirmamos que  $\{J_{\tau_x}^p : p \geq 0\}$  es un proceso de Galton–Watson con distribución de reproducción

$$v(k) = \int_0^\infty \frac{(r\Pi(\mathbb{R}_+)/\alpha)^k}{k!} \exp\left\{-\frac{r}{\alpha}\Pi(\mathbb{R}_+)\right\} \frac{d\Pi(r)}{\Pi(\mathbb{R}_+)}.$$

En efecto, argumentando como en la demostración de la proposición 2, condicional en  $J_{\tau_x}^0$ ,  $\{J_{\tau_x}^p : p \geq 1\}$  tiene la misma distribución que la suma de  $J_{\tau_x}^0$  copias independientes de  $\{J_{\tau_x}^{p-1} : p \geq 1\}$ . La afirmación se sigue al notar que la distribución de  $J_{\tau_x}^0$  es precisamente  $v$ . Observemos que la media de  $v$  satisface

$$\sum_{k=0}^\infty kv(k) = \frac{1}{\alpha} \int_0^\infty r d\Pi(r),$$

lo que nos dice que el proceso  $J_{\tau_x}$  es crítico o subcrítico acorde a si  $X$  es recurrente o deriva a  $-\infty$  de forma respectiva.

La segunda observación que haremos será bajo el supuesto  $\Pi(\mathbb{R}_+) = \infty$ . En este caso se tendrá, para cada  $p \geq 1$ ,

$$\mathbb{P}(Z_{\tau_x}^p > 0) = \lim_{\lambda \rightarrow \infty} \mathbb{E}[1 - \exp\{-\lambda Z_{\tau_x}^p\}] = 1,$$

situación que no ocurre cuando  $\Pi(\mathbb{R}_+) < \infty$ . Esta propiedad está relacionada a que el conjunto  $\{t : N_t = \infty\}$  es distinto del vacío. Ciertamente, podemos construir una sucesión aleatoria  $\{t_n\}$  que cumpla  $\Delta X_{t_n} > 0$  y  $t_n < t_{n+1} < \inf\{s > t_n : X_s = X_{t_n}\}$ . Para esta sucesión  $N_{t_n} \geq n$  y entonces  $t := \lim_n t_n$  satisface  $N_t = \infty$ . De hecho este argumento prueba algo más fuerte, que  $\{t : N_t = \infty\}$  es denso en  $\mathbb{R}_+$ .

### 3.4. El proceso $\rho$

El objetivo de esta sección será investigar al proceso  $\rho = \{\rho_t : t \geq 0\}$  definido por la relación

$$\rho_t := \{\rho_t(s) : s \in \mathcal{J}_t\}.$$

Este proceso nos dice, para cada tiempo  $t \geq 0$ , el estado de la cola al tiempo  $t$ . Debido al lema 3.1, el conjunto  $\mathcal{J}_t$  se puede escribir como una sucesión creciente  $s_t^1 < s_t^2 < \dots$ , la cual puede ser finita o infinita. Puesto que, casi seguramente,  $\sum_{s \in \mathcal{J}_t} \rho_t(s) = X_t - I_t < \infty$  para cada  $t \geq 0$ , podemos ver a  $\rho$  como un proceso aleatorio con valores en  $\ell^1$ , considerando  $\rho_t^k = \rho_t(s_t^k)$  si  $k \leq N_t$  y  $\rho_t^k = 0$  en otro caso. En particular tendremos  $\rho_t = 0$  si, solamente si,  $N_t = 0$ . Más aún, a pesar de la observación con la que se cerró la sección 3.3, el lema 3.1 muestra que, para cada  $t$  fijo,  $\rho_t \in \ell_0^1$  casi seguramente, donde  $\ell_0^1$  es el subconjunto de  $\ell^1$  de sucesiones con soporte finito.

Antes de presentar el resultado principal de la sección, introducimos notación que usaremos. Denotaremos por  $\ell_+^1$  a la clase de elementos de  $\ell^1$  con entradas no negativas. Para  $a \in \ell_+^1$  definiremos a la “longitud” de  $a$  mediante

$$\|a\|_1 = \sum_{j=1}^{\infty} a^j.$$

Dicho esto, para  $a \in \ell_+^1$  y  $h > 0$  definimos a  $k_h a$  como el elemento de  $\ell_0^1$  cuya longitud es  $(\|a\|_1 - h)_+$  y que coincide con  $a$  antes de la última entrada distinta de cero de  $k_h a$ . A saber,  $k_h a = 0$  si  $\|a\|_1 \leq h$  y si  $\|a\|_1 > h$  podemos encontrar un único entero  $m$  tal que

$$\sum_{i=m+1}^{\infty} a^i \leq h < \sum_{i=m}^{\infty} a^i,$$

y entonces ponemos

$$(k_h a)^i = \begin{cases} a^i & \text{si } i < m, \\ \sum_{i=m}^{\infty} a^i - h & \text{si } i = m, \\ 0 & \text{si } i > m, \end{cases}$$

de tal forma que  $\|k_h a\|_1 = (\|a\|_1 - h)_+$ . Por construcción  $k_h a \in \ell_0^1$ . Es posible ver

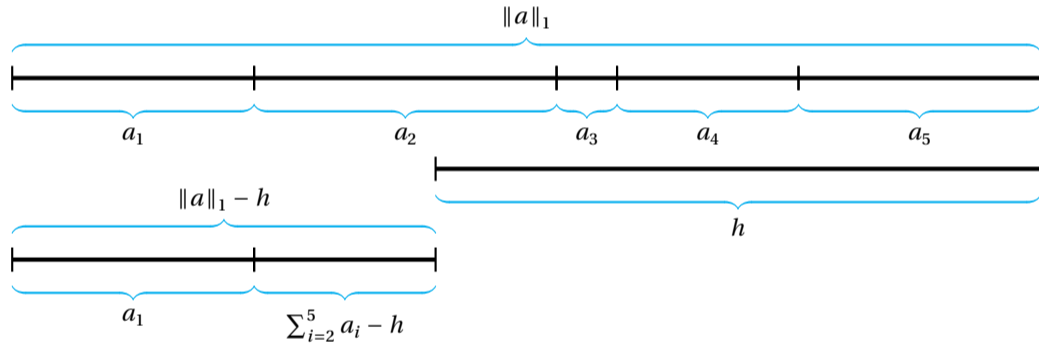


Figura 2: Visualización de la acción de  $k_h$ .

que

$$m = \inf \left\{ k \in \mathbb{N}_0 : \sum_{i=k+1}^{\infty} a^i \leq h \right\}.$$

Además de lo anterior, dados  $a \in \ell_+^1 \cap \ell_0^1$  y  $b \in \ell_+^1$  denotamos por  $[a, b]$  al elemento de  $\ell_+^1$  que se obtiene de concatenar  $a$  y  $b$ , esto es: si  $m = \sup\{i : a^i > 0\}$ ,

$$[a, b]^i = \begin{cases} a^i & \text{si } i \leq m, \\ b^{i-m} & \text{si } i > m. \end{cases}$$

El cometido de la sección se resume en el siguiente resultado, que probaremos en lo que resta del apartado. Denotamos por  $\{\mathcal{F}_t : t \geq 0\}$  a la filtración canónica de  $X$  aumentada por los conjuntos nulos de  $\mathcal{F}_\infty$ .

### 1. Proposición

*El proceso  $\rho$  es un proceso càdlàg, fuertemente markoviano con respecto a la filtración*

$\{\mathcal{F}_t : t \geq 0\}$ , para el cual 0 es un punto recurrente. Una medida invariante para  $\rho$  se caracteriza mediante

$$\begin{aligned} \nu(\Phi) &= \Phi(0) + N\left(\int_0^{T_0} \Phi(\rho_s) ds\right) \\ &= \Phi(0) + \sum_{k=1}^{\infty} R^k \int \Phi(x_1, \dots, x_k, 0, 0, \dots) \tilde{\Pi}^{\otimes k}(dx_1, \dots, dx_k), \end{aligned}$$

donde  $\Phi : \ell_+^1 \rightarrow \mathbb{R}_+$  es medible y

$$d\tilde{\Pi}(r) = \frac{\Pi([r, \infty))}{\int_0^\infty s d\Pi(s)} dr \quad \gamma \quad R = \frac{1}{\alpha} \int_0^\infty r d\Pi(r).$$

Demostración: Para  $s, t \geq 0$  definamos  $X_s^{(t)} := X_{t+s} - X_t$  e  $I_s^{(t)} := \inf_{0 \leq r \leq s} X_r^{(t)}$ . Con probabilidad uno, para cada  $s > 0$  y  $t \geq 0$  se tendrá  $I_s^{(t)} < 0$ . Sea  $\rho^{(t)} = \{\rho_s^{(t)} : s \geq 0\}$  el proceso análogo a  $\rho$  para  $X^{(t)} = \{X_s^{(t)} : s \geq 0\}$ . Entonces, casi seguramente, para cualquier elección de  $t \geq 0$  y  $s \geq 0$ ,  $\rho_s^{(t)} \in \ell_+^1$ . En vista de la interpretación de teoría de colas, con probabilidad uno, para cualesquiera  $t \geq 0$  y  $s > 0$  se cumplirá

$$(1) \quad \rho_{t+s} = [k_{-I_s^{(t)}} \rho_t, \rho_s^{(t)}].$$

La continuidad por la derecha de  $\rho$  se sigue de inmediato, al observar que casi seguramente, para  $t \geq 0$ ,

$$\lim_{s \rightarrow 0^+} I_s^{(t)} = 0 \quad \text{y} \quad \lim_{s \rightarrow 0^+} \|\rho_s^{(t)}\|_1 = 0.$$

Notando que para  $s < t$ ,

$$\rho_t = [k_{-I_{t-s}^{(s)}} \rho_s, \rho_{t-s}^{(s)}]$$

deducimos, mediante argumentos análogos a los dados, que  $\rho$  es càdlàg. Más aún, de esta representación vemos que  $\rho$  y  $X$  comparten los tiempos de discontinuidades y que  $\rho_t = [\rho_{t-}, \Delta X_t e_1]$ , donde  $e_1^i = \delta_{1i}$ .

Usando la igualdad (1) definimos al proceso  $\rho$  comenzado en el elemento  $a \in \ell_+^1$  mediante

$$\rho_t^a = [k_{-I_t} a, \rho_t].$$

Luego, la propiedad fuerte de Markov de  $\rho$  se sigue de la misma propiedad de  $X$  y (1).

Usando el hecho de que 0 es regular y recurrente para  $X - I$  se sigue inmediatamente que 0 es un punto regular y recurrente para  $\rho$ . Adicionalmente, la ley de  $\rho$  bajo  $N$  se puede identificar con la medida de excursiones de  $\rho$  fuera de 0. Entonces, una medida invariante para  $\rho$  será

$$\nu(B) = \mathbb{1}_B(0) + N\left(\int_0^{T_0} \mathbb{1}_B(\rho_t) dt\right).$$

Debido a la igualdad (1.4),

$$\nu(B) = \mathbb{1}_B(0) + \int_0^\infty \mathbb{E}_r^0 \left[ \int_0^{T_0} \mathbb{1}_B(\rho_t) dt \right] d\Pi(r).$$

El lema 3.1 implica que  $\nu$  tiene soporte en  $\ell_0^1$ . Dado  $k \in \mathbb{N}$ , identificamos  $B \in \mathcal{B}((0, \infty)^k)$  con el subconjunto de  $\ell_0^1$  dado por  $\{[a, 0] : a \in B\}$ . Buscando calcular  $\nu(B)$ , denotemos por  $\tilde{B}$  al conjunto de trayectorias  $\omega : [0, t] \rightarrow \mathbb{R}$  cuyo máximo tenga  $k$  saltos sucesivos  $\alpha_1, \dots, \alpha_k$  de tal forma que  $(\alpha_k, \dots, \alpha_1) \in B$ . Sean  $\gamma_1, \gamma_2, \dots$  los tiempos donde  $S$  crece. Entonces

$$\begin{aligned} \nu(B) &= \int_0^\infty \mathbb{E}_r^0 \left[ \int_0^{T_0} \mathbb{1}_B(\rho_t) dt \right] d\Pi(r) \int_0^\infty \mathbb{E}_r^0 \left[ \int_0^{T_0} \mathbb{1}_{\tilde{B}}(\hat{X}^{(t)}) dt \right] d\Pi(r) \\ &= \int_0^\infty \int_0^\infty \mathbb{E}_r^0 [\mathbb{1}_{\tilde{B}}(\hat{X}^{(t)}) \mathbb{1}_{\{t < T_0\}}] dt d\Pi(r). \end{aligned}$$

Luego, notando que bajo  $\mathbb{P}_r^0$ ,  $t < T_0$  si y sólo si  $\inf_{0 \leq s \leq t} X_s > 0$ , situación que es equivalente a  $r > \sup_{0 \leq s \leq t} (X_t - X_s) - (X_t - r)$ , deducimos

$$\begin{aligned} &\int_0^\infty \int_0^\infty \mathbb{E}_r^0 [\mathbb{1}_{\tilde{B}}(\hat{X}^{(t)}) \mathbb{1}_{\{t < T_0\}}] dt d\Pi(r) \\ &= \int_0^\infty \int_0^\infty \mathbb{E} [\mathbb{1}_{\tilde{B}}([X_{[0,t)}, X_t + r]) \mathbb{1}_{\{S_t - X_t < r\}}] dt d\Pi(r) \\ &= \int_0^\infty \int_0^\infty \mathbb{E} [\mathbb{1}_{\{\gamma_{k-1} < t < \gamma_k\}} \mathbb{1}_B(X_t + r - S_t, X_{\gamma_{k-1}} - X_{\gamma_{k-2}}, \dots, X_{\gamma_1}) \mathbb{1}_{\{S_t - X_t < r\}}] dt d\Pi(r) \\ &= \int_0^\infty \mathbb{E} \left[ \mathbb{1}_{\{\gamma_{k-1} < \infty\}} \int_{\gamma_{k-1}}^{\gamma_k} \mathbb{1}_B(X_t + r - S_t, X_{\gamma_{k-1}} - X_{\gamma_{k-2}}, \dots, X_{\gamma_1}) \mathbb{1}_{\{S_t - X_t < r\}} dt \right] d\Pi(r), \end{aligned}$$

en donde se ha usado la notación  $[X_{[0,t)}, X_t + r]$  para denotar la trayectoria càdlàg en  $[0, t]$  que coincide con  $X_s$  para  $s < t$  y es igual a  $X_t + r$  para  $s = t$ . Para entender el cambio de variable realizado en la primera igualdad conviene ver la figura 3, donde se marca el dual con el supremo correspondiente. Luego, al usar la igualdad en distribución  $\hat{X}^{(t)} \stackrel{d}{=} \{X_s : 0 \leq s \leq t\}$ , debemos forzar a que haya un salto en el tiempo  $t$ , el cual es de tamaño  $r$ , lo cual nos obliga a considerar

$$[X_{[0,t)}, X_t + r]_s = \begin{cases} X_s & \text{si } s \in [0, t), \\ X_t + r & \text{si } s = t. \end{cases}$$

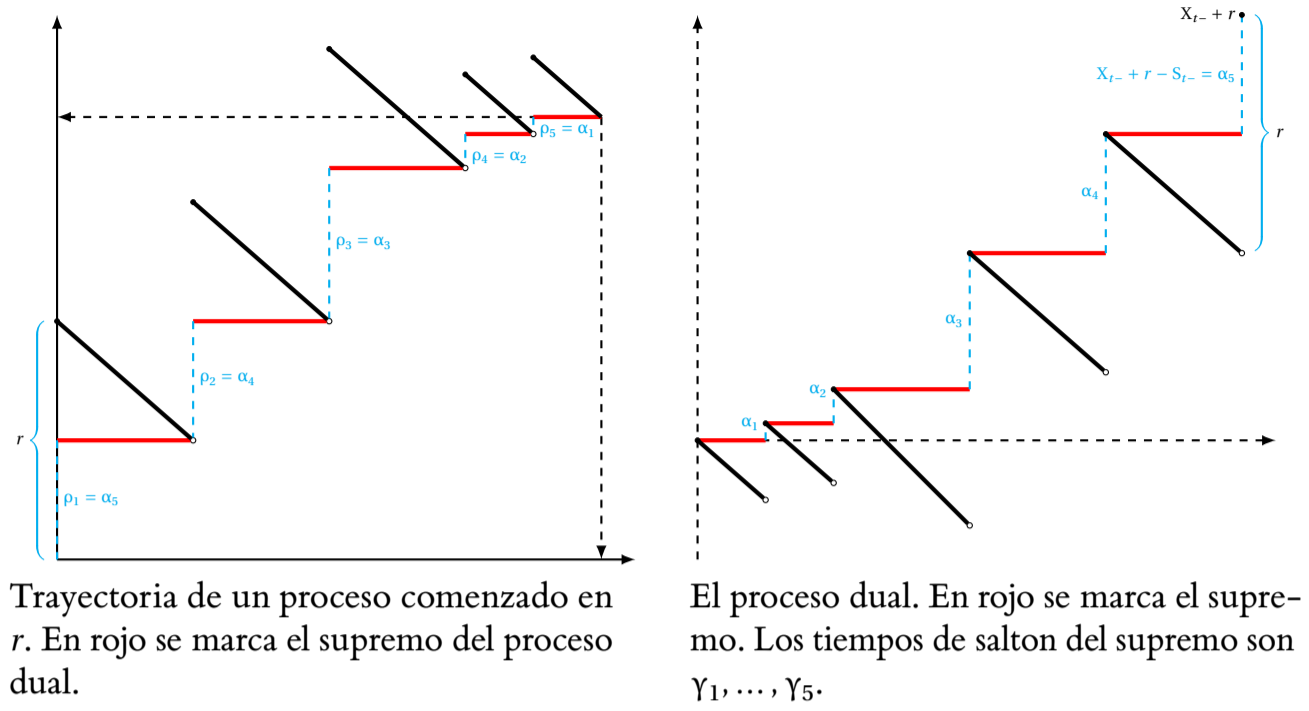
Puesto que el proceso tiene una cantidad contable de saltos, bajo la integral de Lebesgue es equivalente considerar la cantidad anterior con  $X_t$  en lugar de  $X_{t-}$ .

Condicional en el evento  $\{\gamma_{k-1} < \infty\}$ , las variables aleatorias  $X_{\gamma_1}, \dots, X_{\gamma_{k-1}} - X_{\gamma_{k-2}}$  son independientes e idénticamente distribuidas. En vista de (1.1), su distribución en común es  $\tilde{\Pi}$ . Además  $\mathbb{P}(\gamma_{k-1} < \infty) = \mathbb{P}(\gamma_1 < \infty)^{k-1} = R^{k-1}$ . De estas observaciones y el uso de la propiedad fuerte de Markov en  $\gamma_{k-1}$ ,

$$\begin{aligned} &\mathbb{E} \left[ \mathbb{1}_{\{\gamma_{k-1} < \infty\}} \int_{\gamma_{k-1}}^{\gamma_k} \mathbb{1}_B(X_t + r - S_t, X_{\gamma_{k-1}} - X_{\gamma_{k-2}}, \dots, X_{\gamma_1}) \mathbb{1}_{\{S_t - X_t < r\}} dt \right] \\ &= R^{k-1} \int_{(0, \infty)^{k-1}} \mathbb{E} \left[ \int_0^Y \mathbb{1}_B(X_t + r, x_2, \dots, x_k) \mathbb{1}_{\{X_t > -r\}} dt \right] \tilde{\Pi}^{\otimes(k-1)}(dx_2, \dots, dx_k). \end{aligned}$$

Ahora por la última igualdad en (1.3) con  $f(y, r) = \mathbb{1}_B(y, x_2, \dots, x_k)$  obtenemos la primera igualdad en

$$\begin{aligned} &\int_0^\infty \mathbb{E} \left[ \int_0^Y \mathbb{1}_B(X_t + r, x_2, \dots, x_k) \mathbb{1}_{\{X_t > -r\}} dt \right] d\Pi(r) \\ &= \frac{1}{\alpha} \int_{-\infty}^0 \int_{-y}^\infty \mathbb{1}_B(y + r, x_2, \dots, x_k) d\Pi(r) dy \\ &= \frac{1}{\alpha} \int_0^\infty \int_{-r}^0 \mathbb{1}_B(y + r, x_2, \dots, x_k) dy d\Pi(r), \end{aligned}$$

Figura 3: Representación gráfica del proceso  $\rho$  y el proceso dual.

y la segunda igualdad dándose por Fubini. Ahora, considerando el cambio de variable  $x_1 = y + r$ , Fubini y la definición de  $R$  obtenemos

$$\begin{aligned} & \frac{1}{\alpha} \int_0^\infty \int_{-r}^0 \mathbb{1}_B(y + r, x_2, \dots, x_k) dy d\Pi(r) \\ &= \frac{R}{\int_0^\infty s d\Pi(s)} \int_0^\infty \int_{x_1}^\infty \mathbb{1}_B(x_1, \dots, x_k) d\Pi(r) dx_1 \\ &= R \int_0^\infty \mathbb{1}_B(x_1, \dots, x_k) d\tilde{\Pi}(x_1). \end{aligned}$$

Por consiguiente, hilando las igualdades desarrolladas,

$$v(B) = R^k \int_{(0, \infty)^k} \mathbb{1}_B(x_1, \dots, x_k) \tilde{\Pi}^{\otimes k}(dx_1, \dots, dx_k),$$

siendo esto lo que se quería probar. ■

### 3.5. Propiedades de dualidad

El objetivo de esta sección es estudiar el proceso dual de  $\rho$  con respecto a la medida  $v$  dada en la proposición 4.1. Este será un proceso  $\eta$  que también se puede interpretar en términos del sistema de colas. De la misma manera que  $\rho_t$  representa el servicio remanente de los clientes en el sistema al tiempo  $t$ ,  $\eta_t$  representará el servicio que ya se le haya dedicado a dichos clientes. Precizando lo anterior, pondremos  $\eta_t = (\eta_t^1, \eta_t^2, \dots)$ , con

$$\eta_t^k = \begin{cases} \Delta X_{s_t^k} - \rho_t^k & \text{si } k \leq N_t, \\ 0 & \text{si } k > N_t, \end{cases}$$

donde  $N_t$  se entiende como en el lema 3.1.

Puesto que  $\sum_{0 \leq s \leq t} \Delta X_s < \infty$  casi seguramente, con probabilidad uno,  $\eta$  tomará valores en  $\ell_+^1$ . Más aún, para cada  $t \geq 0$ ,  $\mathbb{P}(\eta_t \in \ell_+^1) = 1$ .

A continuación enunciaremos el resultado principal de la sección.

**1. Proposición**

El par  $(\rho, \eta)$  es un proceso fuertemente markoviano càdlàg con valores en  $\ell_+^1 \times \ell_+^1$ . Una medida invariante para  $(\rho, \eta)$  queda caracterizada mediante

$$\Theta(\Phi) = \sum_{k=1}^{\infty} R^k \int \Phi((x_1, \dots, x_k, 0, \dots), (y_1, \dots, y_k, 0, \dots)) \theta^{\otimes k}(dx_1 dy_1, \dots, dx_k dy_k) + \Phi(0)$$

donde la medida  $\theta$  queda definida sobre  $(0, \infty)^2$  por

$$\int f(x, y) \theta(dx dy) = \frac{1}{\int_0^\infty z d\Pi(z)} \int_0^\infty \int_0^z f(x, z-x) dx d\Pi(z).$$

El proceso dual de  $(\rho, \eta)$  con respecto a  $\Theta$  es  $(\eta, \rho)$ .

Antes de demostrar la proposición 1, se enunciará y probará un resultado auxiliar.

**2. Proposición**

Los procesos  $\{(\rho_{(T_0-t)-}, \eta_{(T_0-t)-}) : 0 \leq t \leq T_0\}$  y  $\{(\eta_t, \rho_t) : 0 \leq t \leq T_0\}$  tienen la misma distribución bajo  $N$ .

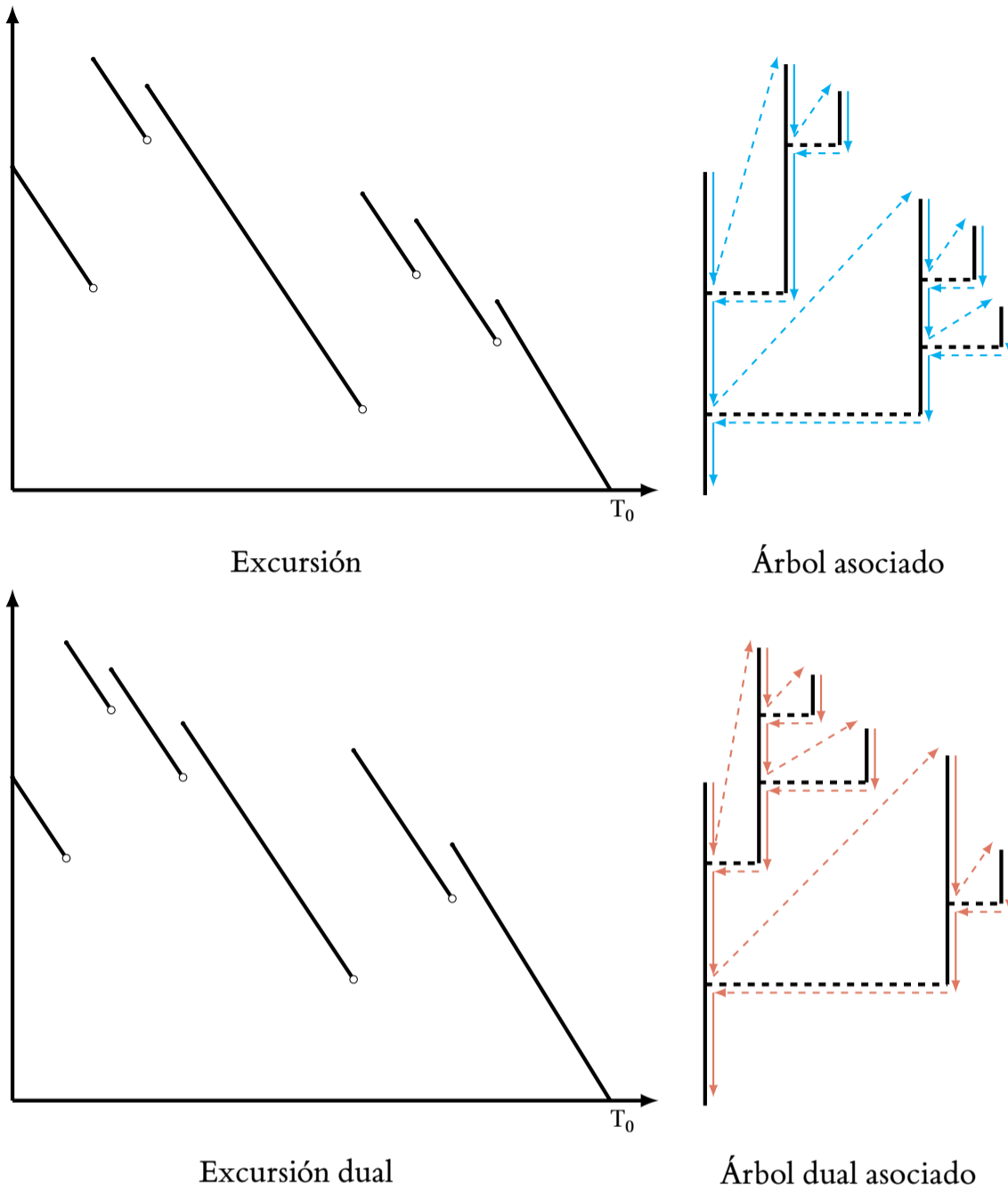
Demostración: Comenzamos haciendo notar que podemos restringirnos al caso en el cual  $\Pi$  es finita. Ciertamente, si  $\rho^\varepsilon$  y  $\eta^\varepsilon$  son los procesos análogos a  $\rho$  y  $\eta$  bajo  $X^\varepsilon$ , entonces procedemos como sigue: observamos que el enunciado es equivalente a que los procesos  $\{(\rho_{(\tau_x-s)-}, \eta_{(\tau_x-s)-}) : 0 \leq s \leq \tau_x\}$  y  $\{(\eta_s, \rho_s) : 0 \leq s \leq \tau_x\}$  tienen la misma distribución bajo  $\mathbb{P}_0$ . Entonces, si el resultado fuese cierto cuando  $\Pi(\mathbb{R}_+) < \infty$ , los procesos  $\{(\rho_{(\tau_x^\varepsilon-s)-}^\varepsilon, \eta_{(\tau_x^\varepsilon-s)-}^\varepsilon) : 0 \leq s \leq \tau_x^\varepsilon\}$  y  $\{(\eta_s^\varepsilon, \rho_s^\varepsilon) : 0 \leq s \leq \tau_x^\varepsilon\}$  tendrán la misma distribución. El caso general se seguirá al tomar el límite conforme  $\varepsilon \rightarrow 0$ , notando que  $\tau_x^\varepsilon \rightarrow \tau_x$ ,  $\rho_t^\varepsilon \rightarrow \rho_t$  y  $\eta_t^\varepsilon \rightarrow \eta_t$  para cada  $t \geq 0$ , casi seguramente con respecto a  $\mathbb{P}_0$ .

Entonces probaremos el resultado bajo el supuesto  $\Pi(\mathbb{R}_+) < \infty$ . Para ello usaremos una representación de las excursiones de  $X - I$  como un árbol marcado, que describiremos usando la interpretación como sistema de colas. Cada cliente, o salto de  $X$ , se representa con un vértice en el árbol. El cliente inicial corresponde a la raíz del árbol. Los hijos de cada cliente son aquellos clientes que interrumpen su servicio. A cada vértice se le asignan dos marcas:

- La primera es el tiempo de servicio para el cliente, esto es el tamaño del salto.
- La segunda, definida salvo para la raíz, es el tiempo de servicio que se le ha dedicado a “la madre” antes de que un cliente lo interrumpa.

Así, podemos pensar en la primera marca como una línea vertical, cuya longitud es el tamaño del salto, dividida en subintervalos que corresponden a tiempos de servicio sin interrupciones. Cada punto que divide este intervalo corresponde exactamente a un hijo de dicho cliente. En este caso, la segunda marca del hijo es la distancia entre tal punto y el principio del segmento de la madre (situado en la parte superior). En la figura 4 se muestra el árbol asociado a una trayectoria  $\{X_t : 0 \leq t \leq T_0\}$  bajo  $N$ .

Para especificar la distribución del árbol bajo la medida de excursión  $N$ , hacemos notar que las longitudes de los segmentos son independientes e idénticamente distribuidos, con distribución común  $\Pi/\Pi(\mathbb{R}_+)$ . Aunado a lo anterior, para un segmento dado, los puntos que lo dividen en subintervalos se distribuyen acorde a una me-

Figura 4: Excursión de  $X - I$  y árbol asociado junto a los duales.

dida aleatoria de Poisson con intensidad  $\alpha^{-1}\Pi(\mathbb{R}_+) dx$ , la cual es independiente del tamaño del segmento e independiente de segmentos diferentes.

Del árbol es posible recuperar los procesos  $\rho$  y  $\eta$ . En efecto, basta considerar una partícula que recorre los segmentos a una velocidad constante  $\alpha$ , la cual comienza en la parte superior del segmento raíz. Cuando la partícula llega a un punto de división, salta al inicio del segmento “hijo”. Asimismo, cuando la partícula finaliza el recorrido de un segmento, regresa al punto de donde partió en el segmento “madre”. La partícula se detiene al llegar a la parte inferior del segmento raíz. Esta dinámica se muestra en la figura 4. Con la dinámica descrita,  $\rho_t$  se obtiene al considerar la longitud de la parte que la partícula no ha visitado de cada uno de los segmentos ancestros del segmento que se encuentra recorriendo al tiempo  $t$ . En cuando a  $\eta_t$  debemos considerar la longitud de las partes ya visitadas al tiempo  $t$ .

Al invertir el tiempo en el cual se mueve la partícula, ahora comenzaremos en la parte inferior de la raíz. Un ejemplo de la inversión del tiempo, y el árbol que se obtiene, se muestra en la figura 4. Usando la interpretación de  $\eta_t$  y  $\rho_t$  dada en el párrafo previo, el efecto de la inversión del tiempo es reemplazar  $\{(\eta_t, \rho_t) : 0 \leq t \leq T_0\}$

por  $\{(\rho_{(T_0-t)-}, \eta_{(T_0-t)-}) : 0 \leq t \leq T_0\}$ . La estructura aleatoria del árbol construido muestra que ambos procesos tienen la misma distribución. ■

Demostración de la proposición 1: Comenzamos extendiendo la expresión (4.1) a  $(\rho, \eta)$ . Para  $h > 0$  definimos a la función  $\tilde{k}_h : \ell_+^1 \times \ell_+^1 \rightarrow \ell_+^1 \times \ell_+^1$ , dada por  $(a, b) \mapsto (a', b')$  como sigue: Si  $\|a\|_1 \leq h$ , entonces  $a' = b' = 0$ . En otro caso, es decir si  $\|a\|_1 > h$ , ponemos  $a' = k_h a$  y

$$b'^i = \begin{cases} b^i & \text{si } i < m, \\ b^m + \left( \sum_{i=m}^{\infty} a^i - h \right) & \text{si } i = m, \\ 0 & \text{si } i > m, \end{cases}$$

donde  $m \geq 1$  es la de la definición de  $k_h a$ . Entonces, usando la interpretación en términos de colas, obtenemos que con probabilidad uno se cumple

$$(\rho_{t+s}, \eta_{t+s}) = \left[ \tilde{k}_{-1_s^{(t)}}(\rho_t, \eta_t), (\rho_s^{(t)}, \eta_s^{(t)}) \right] \quad \text{para cualesquiera } t \geq 0 \text{ y } s > 0,$$

donde la concatenación se realiza de forma separada para  $\rho$  y  $\eta$ . Luego, la propiedad fuerte de Markov de  $(\rho, \eta)$ , así como la propiedad càdlàg de las trayectorias, se deduce como en la demostración de la proposición 4.1.

De igual manera, el cálculo de la medida invariante es similar a aquel realizado en la demostración de la proposición 4.1, usando que la medida de excursión de  $(\rho, \eta)$  fuera de cero es la ley de  $(\rho, \eta)$  bajo  $N$ . En efecto, siguiendo las mismas ideas, deducimos la igualdad

$$\begin{aligned} \nu(A \times B) &= \int_0^\infty \mathbb{E} \left[ \mathbb{1}_{\{\gamma_{k-1} < \infty\}} \int_{\gamma_{k-1}}^{\gamma_k} \mathbb{1}_A(X_t + r - S_t, X_{\gamma_{k-1}} - X_{\gamma_{k-2}}, \dots, X_{\gamma_1}) \right. \\ &\quad \left. \times \mathbb{1}_B(S_t - X_t, X_{\gamma_{k-2}} - X_{\gamma_{k-1}-}, \dots, -X_{\gamma_1-}) dt \right] d\Pi(r). \end{aligned}$$

Luego, se verifica fácilmente que, condicional en  $\{\gamma_{k-1} < \infty\}$ , los pares de variables aleatorias  $(X_{\gamma_1}, -X_{\gamma_1-}), \dots, (X_{\gamma_{k-1}} - X_{\gamma_{k-2}}, X_{\gamma_{k-2}} - X_{\gamma_{k-1}-})$  son independientes e idénticamente distribuidas con distribución común  $\theta$ . Entonces basta usar la propiedad fuerte de Markov y

$$\begin{aligned} \mathbb{E} \left[ \int_0^\gamma f(X_t + r, -X_t) \mathbb{1}_{\{X_t > -r\}} dt \right] &= \frac{1}{\alpha} \int_{-r}^0 f(y + r, -y) dy \\ &= \frac{R}{\int_0^\infty z d\Pi(z)} \int_0^r f(x, r - x) dx \end{aligned}$$

para obtener la forma dada de la medida invariante.

Finalmente, la afirmación sobre la dualidad se sigue de la proposición 2. ■

### 3.6. Combinando ramificación y movimiento espacial

Para finalizar el capítulo consideraremos la construcción de ciertas cadenas de Markov ramificantes. A lo largo de la sección  $(E, \mathcal{E})$  será un espacio medible auxiliar, el cual se encuentra equipado con un kernel de transición  $Q$ . Denotaremos por  $Q_y$  a la ley de la  $Q$ -cadena de Markov iniciada en  $y \in E$ . El proceso en el espacio canónico  $E^{\mathbb{N}_0}$  se denota por  $Y = \{Y_n : n \in \mathbb{N}_0\}$ .

Dado un punto  $y_0 \in E$  a continuación definimos una extensión  $\{(\rho_t, q_t) : t \geq 0\}$  del proceso de Markov  $\rho$ . Para cada  $t \geq 0$ ,  $q_t$  es una sucesión, a lo más numerable,  $(q_t^0, q_t^1, \dots)$  con valores en  $E$ , donde  $q_t^0 = y_0$ . Si  $N(q_t)$  representa la cardinalidad de la sucesión  $q_t$ , entonces  $N(q_t) = N_t + 1$  para cada  $t \geq 0$ , casi seguramente. En adición el proceso  $q = \{q_t : t \geq 0\}$  es un proceso de Markov (no homogéneo) bajo la distribución condicional dada por  $\rho$ . Sus kernels de transición los describiremos para  $0 \leq s < t$ , considerando  $m_{s,t} = \inf_{s \leq r \leq t} N_r$ . La distribución condicional de  $q_t$ , dado  $q_s$  y  $\rho$  es:

- (i) La distribución de  $(Y_0, \dots, Y_{N_t})$  bajo  $Q_{y_0}$  si  $m_{s,t} = 0$ .
- (ii) La distribución de  $(q_s^0, \dots, q_s^{m_{s,t}}, Y_1, \dots, Y_{N_t - m_{s,t}})$  bajo  $Q_{q_s^{m_{s,t}}}$  si  $m_{s,t} > 0$ .

Es posible construir una versión del proceso  $q$  de tal forma que, con probabilidad uno,  $q_t^i = q_s^i$  para cada  $i \in [m_{s,t}]$  y cualquier elección de  $s < t$ . Tal versión es única salvo indistinguibilidad.

Para cada entero  $k \geq 0$  denotaremos por  $\tilde{Z}_k^x$  al elemento aleatorio de  $\mathcal{M}_f(E)$  definido mediante

$$\tilde{Z}_k^x := \int_0^{\tau_x} \mathbb{1}_{\{N_t=k\}} \delta_{q_t^k} dt.$$

### 1. Proposición

El proceso  $\{\tilde{Z}_k^x : k \in \mathbb{N}_0\}$  es una cadena de Markov en  $\mathcal{M}_f(E)$  que comienza en  $\alpha^{-1}x\delta_{y_0}$ . Su kernel de transición  $\tilde{P}(\zeta, d\xi)$  queda caracterizado por su funcional de Laplace, dado por

$$\int e^{-\langle \xi, \Lambda \rangle} \tilde{P}(\zeta, d\xi) = \exp\left\{-\int_E \int_E \int_0^\infty (1 - e^{-r\Lambda(z)/\alpha}) d\Pi(r) Q(y, dz) d\zeta(y)\right\}$$

para cualquier función medible y no negativa  $\Lambda$ .

# Convergencia de funcionales de árboles de Galton–Watson

# 4

Después de haber estudiado un poco de teoría de excursiones y la conexión entre procesos de Lévy de variación finita que son recurrentes o derivan a  $-\infty$ , retornamos al escenario planteado en el capítulo 1. Ya hemos caracterizado los árboles discretos de Galton–Watson y hemos estudiado la convergencia de las funciones de altura. En el capítulo actual usaremos esta convergencia, dada en el teorema 1.3.2, para obtener otros resultados de convergencia. Asimismo, se dará el resultado análogo para la convergencia de las funciones de contorno.

Antes de comenzar, es pertinente remarcar que las convergencias que se considerarán en el capítulo, no nos dicen que los árboles discretos convergen como tal. Únicamente implican la convergencia de funcionales que los caracterizan. La convergencia de los árboles será un tema a estudiar posteriormente.

## 4.1. Primeras aplicaciones del teorema 1.3.2

Para comenzar, conviene recordar algunas propiedades del movimiento browniano. Denotaremos por  $\beta$  al movimiento browniano estándar iniciado en 0. Entonces existe un proceso continuo y creciente  $L_t = L_t(\beta)$  conocido como el tiempo local de  $\beta$  en 0, ver la sección 2.2 o el apéndice A.1. Debido al teorema 1.1.9 sabemos que si  $N_t^\varepsilon$  cuenta el número de excursiones positivas de  $\beta$  fuera de cero que alcanzan un nivel  $\varepsilon > 0$  y se completan antes del tiempo  $t$ , entonces

$$2\varepsilon N_t^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} L_t$$

para cada  $t \geq 0$ , casi seguramente. En adición, el soporte de la medida  $dL$  coincide, casi seguramente, con el conjunto de ceros  $\mathcal{Z} = \{t \geq 0 : \beta_t = 0\}$ . Además, la identidad de Lévy, teorema 1.2.3, nos dice que

$$(B - \underline{B}, -B) \stackrel{d}{=} (|\beta|, L)$$

donde  $\underline{B}_t = \inf_{0 \leq s \leq t} B_s$ .

Manteniendo la notación de la sección 1.3 definimos, para cada  $n \geq 0$  entera,

$$\Lambda_n = k \quad \text{si, solamente si,} \quad \sum_{i=1}^{k-1} \#\theta_i \leq n < \sum_{i=1}^k \#\theta_i,$$

de tal forma que  $k$  nos dice el índice del árbol a cual pertenece el  $n$ -ésimo vértice que se visita.

Con lo dicho hasta ahora, podemos extender la convergencia establecida en el teorema 1.3.2 a la convergencia

$$(1) \quad \left\{ \left( \frac{1}{\sqrt{p}} H_{\lfloor pt \rfloor}, \frac{1}{\sqrt{p}} \Lambda_{\lfloor pt \rfloor} \right) : t \geq 0 \right\} \xrightarrow{p \rightarrow \infty} \left( \frac{2}{\sigma} |\beta|, \sigma L \right).$$

En efecto, el corolario 1.2.4 y la proposición 1.1.4 implican que

$$\Lambda_n = 1 - \inf_{j \leq n} S_j = 1 - I_n,$$

donde  $\{S_n : n \in \mathbb{N}\}$  es una caminata aleatoria con distribución de saltos descrita en el corolario 1.2.4. En adición, en la sección 1.3 vimos que para cualquier  $A > 0$ ,

$$\sup_{0 \leq t \leq A} \left| \frac{S_{[pt]} - I_{[pt]}}{\sqrt{p}} - \frac{\sigma^2}{2} \frac{H_{[pt]}}{\sqrt{p}} \right| \xrightarrow[p \rightarrow \infty]{\text{c.s.}} 0.$$

Luego, el teorema de Donsker nos da la convergencia

$$\left\{ \left( \frac{1}{\sqrt{p}} H_{[pt]}, \frac{1}{\sqrt{p}} \Lambda_{[pt]} \right) : t \geq 0 \right\} \xrightarrow[p \rightarrow \infty]{} \left( \frac{2}{\sigma} (B - \underline{B}), -\sigma \underline{B} \right),$$

de donde obtenemos (1) gracias a la identidad de Lévy.

Usaremos la convergencia en (1) para estudiar el comportamiento asintótico de un árbol de Galton–Watson condicionado a tener una altura grande. Denotamos por  $h(\theta) := \sup\{|v| : v \in \theta\}$  a la altura de un árbol  $\theta$ . Por otra parte, dados  $x > 0$  y  $p \in \mathbb{N}$  denotamos por  $\theta^{\{x\sqrt{p}\}}$  al árbol aleatorio con distribución

$$\Pi_\mu(dt \mid h(t) \geq x\sqrt{p}),$$

donde recordamos que  $\Pi_\mu$  es la ley del árbol de Galton–Watson con distribución de reproducción  $\mu$ .

Además introducimos la notación  $H^{\{x\sqrt{p}\}}$  para el proceso de altura del árbol  $\theta^{\{x\sqrt{p}\}}$ , con la convención de poner  $H_n^{\{x\sqrt{p}\}}$  siempre que  $n \geq \# \theta^{\{x\sqrt{p}\}}$ .

### 1. Corolario

*Se tiene la convergencia*

$$\left\{ \frac{1}{\sqrt{p}} H_{[pt]}^{\{x\sqrt{p}\}} : t \geq 0 \right\} \xrightarrow[p \rightarrow \infty]{} \frac{2}{\sigma} e^{\sigma x/2},$$

donde  $e^{\sigma x/2}$  es una excursión browniana condicionada a tener altura al menos  $\sigma x/2$ .

La excursión  $e^{\sigma x/2}$  se puede construir de forma explícita. Ciertamente, si ponemos

$$T = \inf \left\{ t \geq 0 : |\beta_t| \geq \frac{\sigma x}{2} \right\},$$

$$G = \sup \{ t \leq T : \beta_t = 0 \}$$

y

$$D = \inf \{ t \geq T : \beta_t = 0 \}$$

entonces tendremos que

$$e_t^{\sigma x/2} = |\beta_{(G+t) \wedge D}| \quad \text{para todo } t \geq 0.$$

Usaremos esta construcción en la demostración.

Demostración: Usando (1), el teorema de representación de Skorokhod nos dice que para cada  $p \geq 1$  podemos encontrar un proceso  $(H^{(p)}, \Lambda^{(p)})$  tal que

$$(H^{(p)}, \Lambda^{(p)}) \stackrel{d}{=} \left\{ \left( \frac{1}{\sqrt{p}} H_{[pt]}, \frac{1}{\sqrt{p}} \Lambda_{[pt]} \right) : t \geq 0 \right\}$$

y, con probabilidad uno,

$$(H^{(p)}, \Lambda^{(p)}) \xrightarrow{p \rightarrow \infty} \left( \frac{2}{\sigma} |\beta|, \sigma L \right)$$

uniformemente en compactos.

Definimos

$$T^{(p)} = \inf\{t \geq 0 : |H_t^{(p)}| \geq x\},$$

$$G^{(p)} = \sup\{t \leq T^{(p)} : H_t^{(p)} = 0\} - \frac{1}{p}$$

y

$$D^{(p)} = \inf\{t \geq T^{(p)} : H_t^{(p)} = 0\}.$$

La razón para restar el término  $p^{-1}$  en la expresión de  $G^{(p)}$  se debe a la parte entera de  $p^{-1/2}H_{[pt]}$  y que queremos que el proceso  $H_{G^{(p)}+t}^{(p)}$  se quede en 0 a lo largo del intervalo  $[0, p^{-1})$ .

Por construcción, el proceso  $\{H_{(G^{(p)}+t) \wedge D^{(p)}} : t \geq 0\}$  tiene la misma distribución que el proceso de altura, debidamente reescalado, del primer árbol  $\theta_1, \theta_2, \dots$  cuya altura sea al menos  $x\sqrt{p}$ , que por definición tiene la misma distribución que  $\theta^{[x\sqrt{p}]}$ . Consecuentemente

$$\{H_{(G^{(p)}+t) \wedge D^{(p)}} : t \geq 0\} \stackrel{d}{=} \left\{ \frac{1}{\sqrt{p}} H_{[pt]}^{[x\sqrt{p}]} : t \geq 0 \right\}$$

y entonces el corolario se obtendrá al demostrar las convergencias casi seguras

$$G^{(p)} \xrightarrow{p \rightarrow \infty} G \quad \text{y} \quad D^{(p)} \xrightarrow{p \rightarrow \infty} D.$$

De la misma definición de  $T$  se sigue inmediatamente que

$$T^{(p)} \xrightarrow[p \rightarrow \infty]{\text{c.s.}} T.$$

Por ello también obtenemos las desigualdades casi seguras

$$\liminf_{p \rightarrow \infty} D^{(p)} \geq D \quad \text{y} \quad \limsup_{p \rightarrow \infty} G^{(p)} \leq G.$$

Demostremos que  $\limsup_p D^{(p)} \leq D$  casi seguramente. Un argumento similar nos dará la desigualdad correspondiente para  $\liminf_p G^{(p)}$ . Fijando  $t > 0$  notamos que con probabilidad uno, en  $\{D < t\}$ ,  $L_D < L_t$  y por ende, en  $\{D < t\}$ ,  $\sigma L_D < \Lambda_t^{(p)}$  para  $p$  suficientemente grande, con probabilidad uno. Ahora, como  $L_T = L_D$ ,  $T^{(p)}$  converge a  $T$  y  $\Lambda_{T^{(p)}}^{(p)}$  converge a  $\sigma L_T$  deducimos que

$$\Lambda_t^{(p)} > \Lambda_{T^{(p)}}^{(p)} \quad \text{para } p \text{ suficientemente grande, casi seguramente en } \{D < t\}.$$

Resta observar que  $\Lambda^{(p)}$  es constante en el intervalo  $[T^{(p)}, D^{(p)})$  para concluir que

$$t \geq D^{(p)} \quad \text{para } p \text{ suficientemente grande, casi seguramente en } \{D < t\}.$$

En particular  $\limsup_p D^{(p)} \leq D + \varepsilon$  para todo  $\varepsilon > 0$ , con lo cual obtenemos que con probabilidad uno se satisface  $\limsup_p D^{(p)} \leq D$ . ■

Al reemplazar  $p$  por  $p^2$  y tomar  $x = 1$  deducimos del corolario que

$$\frac{1}{p^2} \# \theta^{(p)} \xrightarrow[p \rightarrow \infty]{} \zeta_{\sigma/2},$$

donde  $\zeta_{\sigma/2}$  es la longitud de la excursión  $e^{\sigma/2}$ . En efecto, esto es una consecuencia inmediata de la convergencia de  $D^{(p)} - G^{(p)}$  a  $\zeta_{\sigma/2} = D - G$  y que por construcción

$$\# \theta^{(p)} \stackrel{d}{=} p^2(D^{(p^2)} - G^{(p^2)}) + 1$$

usando la notación de la prueba precedente.

La transformada de Laplace de la ley límite es conocida y dada por

$$\mathbb{E}[e^{-\lambda \zeta_{\sigma/2}}] = \frac{\sigma \sqrt{2\lambda}/2}{\sinh(\sigma \sqrt{2\lambda}/2)} \exp\left\{-\frac{\sigma \sqrt{2\lambda}}{2}\right\}.$$

Esta identidad es una consecuencia de la descomposición de Williams de la medida de excursiones de Itô, ver el teorema 31 de las notas de Pardo Millán (s.f.-b), y las fórmulas del tiempo de llegada al nivel  $\sigma/2$  de un proceso de Bessel de dimensión tres comenzado en cero.

Ahora discutiremos las *medidas de ocupación*, para lo cual estaremos interesados en un bosque de árboles finito cuyo tamaño tenderá a infinito junto a  $p$ . De manera más precisa fijamos  $b > 0$  y ponemos

$$H_n^p = \begin{cases} H_n & \text{si } \Lambda_n \leq bp, \\ 0 & \text{si } \Lambda_n > bp \end{cases}$$

de tal forma que  $H^p$  es el proceso de altura de una colección de  $[bp]$  árboles de Galton–Watson independientes. Así, obtenemos de (1) que

$$(2) \quad \left\{ \frac{1}{p} H_{[p^2 t]}^p : t \geq 0 \right\} \xrightarrow[p \rightarrow \infty]{} \left\{ \frac{2}{\sigma} |\beta_{t \wedge \tau_{b/\sigma}}| : t \geq 0 \right\},$$

donde definimos

$$\tau_r := \inf\{t \geq 0 : L_t > r\}$$

para toda  $r > 0$ . Efectivamente, podemos escribir

$$\frac{1}{p} H_{[p^2 t]}^p = \frac{1}{p} H_{[p^2(t \wedge \tau_b^{(p)})]}^p,$$

donde  $\tau_b^{(p)} = p^{-2} \inf\{n \in \mathbb{N}_0 : \Lambda_n > bp\} = \inf\{t \geq 0 : p^{-1} \Lambda_{[pt]} > b\}$ , con lo cual vemos, usando que

$$\left( \left\{ \frac{1}{p} H_{[p^2 t]}^p : t \geq 0 \right\}, \tau_b^{(p)} \right) \xrightarrow[p \rightarrow \infty]{} \left( \frac{2}{\sigma} |\beta|, \tau_{b/\sigma} \right)$$

es cierto por (1), la veracidad de (2).

Al considerar  $b = 1$ , de (2), deducimos que para toda  $x > 0$ ,

$$\mathbb{P}\left(\sup_{1 \leq i \leq p} h(\theta_i) > px\right) \xrightarrow[p \rightarrow \infty]{} \mathbb{P}\left(\sup_{t \leq \tau_{1/\sigma}} \frac{2}{\sigma} |\beta_t| > x\right) = 1 - \exp\left\{-\frac{2}{\sigma^2 x}\right\}.$$

La igualdad es una consecuencia de la teoría de excursiones para el movimiento browniano lineal. En efecto, si  $N_t(A)$  denota el número de excursiones que tocan el conjunto  $A$  al tiempo  $t$ , sabemos que esta variable aleatoria tiene distribución Poisson de parámetro  $t\nu(A)$ , donde en el caso del movimiento browniano,

$$\nu(\mathbb{R} \setminus [-x, x]) = \frac{1}{x},$$

de modo que

$$\begin{aligned} \mathbb{P}\left(\sup_{t \leq \tau_{1/\sigma}} \frac{2}{\sigma} |\beta_t| > x\right) &= 1 - \mathbb{P}\left(\sup_{t \leq \tau_{1/\sigma}} |\beta_t| \leq \frac{\sigma x}{2}\right) \\ &= 1 - \mathbb{P}\left(N_{1/\sigma}\left(\mathbb{R} \setminus \left[-\frac{\sigma x}{2}, \frac{\sigma x}{2}\right]^c\right) = 0\right) \\ &= 1 - \exp\left\{-\frac{1}{\sigma} \frac{2}{\sigma x}\right\} = 1 - \exp\left\{-\frac{2}{\sigma^2 x}\right\}. \end{aligned}$$

Por otra parte,

$$\mathbb{P}\left(\sup_{1 \leq i \leq p} h(\theta_i) > px\right) = 1 - \left(1 - \mathbb{P}(h(\theta) > px)\right)^p.$$

Como corolario obtenemos

$$\mathbb{P}(h(\theta) \geq n) \sim \frac{2}{\sigma^2 n}$$

cuando  $n \rightarrow \infty$ .

Ahora, poniendo  $Z_0^p = 0$  y ponemos, para  $n \in \mathbb{N}$ ,

$$Z_n^p = \sum_{i=1}^p \#\{u \in \theta_i : |u| = n\} = \#\{k \in \mathbb{N}_0 : H_k^p = n\}.$$

Por la proposición 1.2.1 sabemos que  $\{Z_n^p : n \geq 0\}$  es un proceso de Galton–Watson con distribución de reproducción  $\mu$ . Entonces podemos aplicar la aproximación difusiva clásica:

## 2. Teorema

Tenemos

$$\left\{\frac{1}{p} Z_{[pt]}^p : t \geq 0\right\} \xrightarrow{p \rightarrow \infty} X,$$

donde  $X$  es un proceso de difusión con generador infinitesimal  $\mathcal{A} = (\sigma^2 x/2) d^2/dx^2$ , el cual es la única solución fuerte de la ecuación diferencial estocástica

$$\begin{aligned} dX_t &= \sigma \sqrt{X_t} dB_t, \\ X_0 &= 1. \end{aligned}$$

Una demostración de este resultado se puede consultar en el teorema IX.1.3 de Ethier y Kurtz (2005). En adición, es sencillo comprobar que, para  $p \geq 1$  dada, los procesos

$$\{Z_n^p : n \in \mathbb{N}_0\} \quad \text{y} \quad \left\{(Z_n^p)^2 - \sigma^2 \sum_{k=0}^{n-1} Z_k^p : n \in \mathbb{N}_0\right\}$$

son martingalas, sugiriendo que el proceso límite debe ser de la forma descrita en el teorema.

El proceso  $X$  es conocido como *difusión de Feller*. Cuando  $\sigma = 2$ , el proceso también es conocido como el proceso de Bessel cuadrado de dimensión cero, ver Revuz y Yor (1999). Notamos que el proceso se absorbe en cero en tiempo finito.

Para simplificar la notación fijemos  $\mu$  y  $\sigma = 2$ . Tomemos  $q$  funciones  $f_1, \dots, f_q \in C_c(\mathbb{R}_+, \mathbb{R}_+)$ . Como consecuencia de (2) obtenemos

$$\left\{ \int_0^{\tau_1^{(p)}} f_i \left( \frac{1}{p} H_{[p^2 t]}^p \right) dt : i \in [q] \right\} \xrightarrow{p \rightarrow \infty} \left\{ \int_0^{\tau_{1/2}} f_i(|\beta_t|) dt : i \in [q] \right\}.$$

Por otro lado,

$$\int_0^{\tau_1^{(p)}} f_i \left( \frac{1}{p} H_{[p^2 t]}^p \right) dt = \frac{1}{p^2} \sum_{n \geq 0} Z_n^p f_i \left( \frac{n}{p} \right) = \int_0^\infty f_i \left( \frac{\lfloor pa \rfloor}{p} \right) \frac{1}{p} Z_{[pa]}^p da,$$

y entonces, por la aproximación difusiva,

$$\left\{ \int_0^{\tau_{1/2}} f_i(|\beta_t|) dt : i \in [q] \right\} \stackrel{d}{=} \left\{ \int_0^\infty f_i(a) X_a da : i \in [q] \right\}.$$

Dicho de otra manera, la medida de ocupación de  $|\beta|$  en el intervalo de tiempo  $[0, \tau_{1/2}]$ , es decir la medida descrita por

$$f \mapsto \int_0^{\tau_{1/2}} f(|\beta_t|) dt,$$

tiene la misma distribución que la medida  $X_a da$ .

## 4.2. Árboles de Galton–Watson con progenie total fija

Otra aplicación del teorema 1.3.2 es recuperar un resultado de Aldous sobre árboles de Galton–Watson condicionados a tener un número grande, y fijo, de vértices. Así como lo hicimos en la sección 1.3, supondremos que la distribución de reproducción  $\mu$  tiene momentos exponenciales pequeños. Esto no es necesario para el resultado pero simplifica las cuentas. Siguiendo a Le Gall, los argumentos que se presentan se basan en el artículo de Marckert y Mokkadem (2003).

Para  $p \in \mathbb{N}$  denotamos por  $\theta^{(p)}$  a un GWT( $\mu$ ) condicionado al evento  $\{\#\theta = p\}$ . Luego, necesitaremos que  $\mathbb{P}(\#\theta = p) > 0$  para  $p \in \mathbb{N}$ , lo cual será cierto si  $\mu(1) > 0$ . Denotamos por  $\Pi_\mu(dt \mid \#\mathbf{t} = p)$  a la distribución de  $\theta^{(p)}$ . En adición representaremos al proceso de altura de  $\theta^{(p)}$  por  $H^{(p)} = \{H_k^{(p)} : k \in [0, p]\}$ , conviniendo en que  $H_p^{(p)} = 0$ .

Por otro lado es necesario introducir la excursión browniana normalizada  $\mathbf{e} = \{\mathbf{e}_t : t \in [0, 1]\}$ , la cual es la excursión browniana condicionada a tener longitud 1. Por ejemplo, podemos considerar la primera excursión positiva de  $\beta$  cuya longitud sea mayor que 1, escribir a su intervalo de excursión como  $[G, D]$  y definir

$$E_t = \beta_{(G+t) \wedge D} \quad \text{con } t \geq 0,$$

y

$$\mathbf{e}_t = \frac{1}{\sqrt{D-G}} E_{(D-G)t} \quad \text{con } t \in [0, 1].$$

**1. Teorema**

*Se cumple la convergencia*

$$\left\{ \frac{1}{\sqrt{p}} H_{[pt]}^{(p)} : t \in [0, 1] \right\} \xrightarrow{p \rightarrow \infty} \frac{2}{\sigma} e.$$

El teorema 1 es similar al teorema 1.3.2 y el corolario 1.1. No obstante, como la evento condicional se “degenera en el límite”, conforme  $p \rightarrow \infty$ , no es posible usar la misma estructura de prueba que se usó en el corolario 1.1. Antes de presentar la prueba, enunciamos un resultado auxiliar, que puede pensarse como una versión condicional del teorema de Donsker. La prueba de este teorema se puede consultar en el artículo de Kaigh (1976). La demostración del resultado también se encuentra en este texto, pero por su longitud ella se relegó a la sección 4.A.

**2. Lema**

Sea  $T_1 = \inf\{n \in \mathbb{N} : S_n = -1\}$ . La distribución del proceso  $\{S_{[pt]}/\sigma\sqrt{p} : t \in [0, 1]\}$ , bajo la distribución condicional  $\mathbb{P}(\cdot \mid T_1 = p)$  converge, conforme  $p \rightarrow \infty$ , a la ley de una excursión browniana normalizada.

En el enunciado anterior también podemos entender a la variable  $T_1$  como el número de vértices de un  $\text{GWT}(\mu)$ .

Antes de continuar con la demostración de teorema 1, recordaremos la fórmula de Kemperman, ver Pitman (2006), que relaciona la probabilidad de que el tiempo de llegada al nivel  $-k$  se dé al tiempo  $p$ , con la probabilidad de que una caminata aleatoria se encuentre en el nivel  $-k$  al tiempo  $p$ .

**3. Lema (Fórmula de Kemperman)**

Sea  $S = \{S_n : n \geq 0\}$  una caminata aleatoria que comienza en 0 tal que  $S_1$  tiene soporte en  $\{-1, 0, \dots\}$  y  $\mathbb{P}(S_1 = -1) > 0$ . Si  $T_k = \inf\{p \in \mathbb{N} : S_p = -k\}$  entonces

$$\mathbb{P}(T_k = p) = \frac{k}{p} \mathbb{P}(S_p = -k).$$

Demostración: Sea  $x = (x_1, \dots, x_p)$  una sucesión con valores en  $\{-1, 0, \dots\}$  cuya suma es  $-k$  para algún  $k \in [p]$ , es decir  $\sum_{i=1}^p x_i = -k$ . Afirmamos que existen  $k$  índices distintos  $i \in [p]$  tales que  $x^{(i)} := (x_{i+1}, \dots, x_n, x_1, \dots, x_i)$  satisface  $s_l^{(i)} = \sum_{j=1}^l x_j^{(i)} > -k$  para toda  $l < p$ .

En efecto, sea  $i = m$  el menor índice tal que  $s_i = \min_{1 \leq j \leq p} s_j$ , con  $s_i = \sum_{j=1}^i x_j$ . El índice  $m$  satisface la propiedad requerida, de modo que vemos que existe al menos un índice. Al reemplazar  $x$  por  $x^{(m)}$  podemos suponer que  $s_i > -k$  para toda  $i < p$ . Dado este supuesto vemos que  $x^{(i)}$  satisface la condición  $s_j^{(i)}$  para  $i < p$  si y sólo si  $s_i < \inf_{1 \leq j \leq i-1} s_j$ , lo cual implica que para alguna  $l \in [k]$  se cumplen  $s_i = -l$  y  $s_j > -l$  para  $j < i$ .

En otras palabras, los índices  $i$  para los cuales se satisface la propiedad requerida son aquellos donde se toca por primera vez el nivel  $-l$  para cada  $l \in [k]$  y hay exactamente  $k$  de tales índices. De lo anterior se obtiene inmediatamente el resultado. ■

Demostración del teorema 1: Consideremos a  $H$  como en el teorema 1.3.2, es decir el proceso de alturas asociado a una sucesión de  $\text{GWT}(\mu)$  independientes. Más aún,

usaremos la relación entre  $H$  y la caminata aleatoria  $S$  dada por la proposición 1.3.1.

Denotemos por  $T_1$  al número de vértices del primer árbol en la sucesión, es decir

$$T_1 = \inf\{n \in \mathbb{N} : H_n = 0\} = \inf\{n \in \mathbb{N}_0 : S_n = -1\}.$$

Por el lema 3, para cada  $p \in \mathbb{N}$ ,

$$\mathbb{P}(T_1 = p) = \frac{1}{p} \mathbb{P}(S_p = -1).$$

En tanto, un resultado clásico de caminatas aleatorias, el teorema del límite local (Davis & McDonald, 1995), nos dice que

$$\lim_{p \rightarrow \infty} \sqrt{p} \mathbb{P}(S_p = -1) = \frac{1}{\sigma \sqrt{2\pi}},$$

y por consiguiente

$$(1) \quad \mathbb{P}(T_1 = p) \sim \frac{1}{\sigma \sqrt{2\pi p^3}} \quad \text{conforme } p \rightarrow \infty.$$

Debido al supuesto de momentos exponenciales pequeños, siguiendo las ideas de la demostración de la convergencia funcional del teorema 1.3.2, en particular de (1.3.3), observamos que existe  $\varepsilon > 0$  tal que para  $p$  suficientemente grande tengamos

$$\mathbb{P}\left(\sup_{0 \leq t \leq 1} \left| \frac{H_{[pt]}}{\sqrt{p}} - \frac{2}{\sigma^2} \frac{S_{[pt]} - I_{[pt]}}{\sqrt{p}} \right| > \frac{1}{p^{1/8}}\right) < \exp\{-p^\varepsilon\}.$$

En efecto, basta tomar  $A = 1$  y  $p$  suficientemente grande en (1.3.3) para obtener la desigualdad previa. Usando esta desigualdad y la relación de equivalencia asintótica dada en (1) vemos que para  $p$  suficientemente grande también se cumple

$$\mathbb{P}\left(\sup_{0 \leq t \leq 1} \left| \frac{H_{[pt]}}{\sqrt{p}} - \frac{2}{\sigma^2} \frac{S_{[pt]} - I_{[pt]}}{\sqrt{p}} \right| > \frac{1}{p^{1/8}} \mid T_1 = p\right) < \exp\{-p^{\varepsilon'}\}$$

para cualquier  $\varepsilon' < \varepsilon$ . Ciertamente, consideremos  $\varepsilon' < \varepsilon$ . Debido a la relación de equivalencia asintótica deducimos que dado  $\delta > 0$ ,

$$\frac{1}{\sigma \sqrt{2\pi p^3} \mathbb{P}(T_1 = p)} \leq 1 + \delta \quad \text{para } p \text{ suficientemente grande,}$$

de modo que

$$\begin{aligned} & \mathbb{P}\left(\sup_{0 \leq t \leq 1} \left| \frac{H_{[pt]}}{\sqrt{p}} - \frac{2}{\sigma^2} \frac{S_{[pt]} - I_{[pt]}}{\sqrt{p}} \right| > \frac{1}{p^{1/8}} \mid T_1 = p\right) \\ & < (1 + \delta) \sigma \sqrt{2\pi p^3} \exp\{-p^\varepsilon\} \\ & = \exp\{-p^{-\varepsilon'}\} \exp\left\{-p^\varepsilon + p^\varepsilon + \frac{3}{2} \log p + K_1\right\}, \end{aligned}$$

con  $K_1 = \log(1 + \delta) + \log \sigma + \log(2\pi)/2$ . Puesto que  $\varepsilon' < \varepsilon$ ,

$$\lim_{p \rightarrow \infty} \left(-p^\varepsilon + p^{\varepsilon'} + \frac{3}{2} \log p\right) = 0,$$

y por ende, para  $p$  suficientemente grande,

$$\exp\{-p^{\epsilon'}\} \exp\left\{-p^{\epsilon} + p^{\epsilon'} + \frac{3}{2} \log p\right\} \leq \exp\{-p^{\epsilon'}\},$$

con lo cual obtenemos la afirmación.

Entonces, como  $I_n = 0$  para  $0 \leq n < T_1$ , obtenemos que

$$\mathbb{P}\left(\sup_{0 \leq t \leq 1} \left| \frac{H_{\lfloor pt \rfloor}}{\sqrt{p}} - \frac{2}{\sigma^2} \frac{S_{\lfloor pt \rfloor}}{\sqrt{t}} \right| > \frac{1}{p^{1/8}} \mid T_1 = p\right) < \exp\{-p^{\epsilon'}\}$$

también es cierta para  $p$  suficientemente grande. Para concluir, resta notar que  $H^{(p)}$  tiene la misma distribución que  $\{H_k : k \in \llbracket 0, p \rrbracket\}$  bajo  $\mathbb{P}(\cdot \mid T_1 = p)$ , pues entonces el resultado se sigue de la última desigualdad y el lema 2. ■

Como consecuencia del teorema 1 obtenemos, para cada  $x > 0$ ,

$$(2) \quad \lim_{p \rightarrow \infty} \mathbb{P}(h(\theta) > x\sqrt{p} \mid \#\theta = p) = \mathbb{P}\left(\sup_{0 \leq t \leq 1} e_t > \frac{\sigma x}{2}\right).$$

Otras consecuencias combinatorias las elucidamos a continuación.

Para ciertas elecciones de  $\mu$ , la medida  $\Pi_{\mu}(d\mathbf{t} \mid \#\mathbf{t} = p)$  coincide con la distribución uniforme sobre cierta clase de árboles combinatorios con  $p$  vértices, de modo que el teorema 1 nos da información sobre la proporción de árboles en tal clase que satisfacen algunas propiedades. Para precisar lo anterior, supongamos primero que  $\mu$  es la distribución geométrica con parámetro  $1/2$ . Entonces, gracias a la proposición 1.2.2 vemos que  $\Pi_{\mu}(d\mathbf{t} \mid \#\mathbf{t} = p)$  es la distribución uniforme sobre  $\mathbf{T}_p$ , el conjunto de todos los árboles enraizados con  $p$  vértices. Con ello, (2) muestra que la altura de un árbol elegido al azar en  $\mathbf{T}_p$  es de orden  $\sqrt{p}$  y, de hecho, nos dice la proporción asintótica de árboles en  $\mathbf{T}_p$  con altura mayor a  $x\sqrt{p}$ . De manera similar, es posible aplicar argumentos análogos a funcionales distintos a la altura. Por ejemplo, si  $0 \leq a < b$  son dados, es posible derivar resultados asintóticos para el número de vértices en la generación  $\lfloor a\sqrt{p} \rfloor$  que tienen descendientes en la generación  $\lfloor b\sqrt{p} \rfloor$ . Sin embargo, las distribuciones límites que se obtienen se encuentran en términos de la excursión browniana normalizada y no siempre son sencillas de calcular de forma explícita.

Continuando con las propiedades combinatorias, si consideramos que  $\mu$  sea la distribución Poisson, entonces podemos obtener la distribución uniforme sobre el conjunto de árboles de Cayley con  $p$  vértices a partir de  $\Pi_{\mu}(d\mathbf{t} \mid \#\mathbf{t} = p)$ . Para mayor completud, recordamos que un *árbol de Cayley*, con  $p$  vértices, es un árbol no ordenado con vértices etiquetados por  $1, 2, \dots, p$ . La raíz puede ser cualquiera de los  $p$  vértices. Una fórmula famosa, debida a Cayley y cuya demostración se puede consultar en el capítulo 33 del libro de Aigner y Ziegler (2018), nos dice que el número de árboles de Cayley con  $p$  vértices es  $p^{p-2}$ . Por consiguiente el número de árboles enraizados de Cayley es  $p^{p-1}$ . Así, si comenzamos con un árbol enraizado distribuido acorde a  $\Pi_{\mu}(d\mathbf{t} \mid \#\mathbf{t} = p)$ , asignamos las etiquetas  $1, 2, \dots, p$  de manera uniforme a los vértices, y olvidamos el orden del árbol con el cual comenzamos, obtenemos un árbol aleatorio con distribución uniforme sobre la clase de árboles enraizados de Cayley con  $p$  vértices. Por ende, el teorema 1 nos da información sobre la propiedad de árboles de Cayley grandes.

Como último ejemplo, tomamos  $\mu = 2^{-1}(\delta_0 + \delta_2)$ , de tal forma que si  $p$  es impar,  $\Pi_\mu(d\mathbf{t} \mid \#\mathbf{t} = p)$  es la distribución uniforme sobre el conjunto de árboles binarios y completos con  $p$  vértices. En principio, no es posible aplicar el teorema 1 pues para la demostración habíamos supuesto que  $\mu(1) > 0$ . No obstante, si nos restringimos a valores impares de  $p$ , se cumplirá  $\mathbb{P}(\#\theta = p) > 0$  y entonces la convergencia del teorema 1 sigue siendo cierta.

A primera vista, parece increíble que estas clases de árboles combinatorios tengan el mismo límite de escala; en particular que la distribución límite que aparece en (2) es la misma en cada clase salvo el valor de la constante  $\sigma$ . Para los árboles binarios y los árboles de Cayley  $\sigma^2 = 1$ , mientras que  $\sigma^2 = 2$  para árboles enraizados y ordenados.

### 4.3. Convergencia de las funciones de contorno

El objetivo de este apartado es explicar de forma sucinta cómo es que podemos enunciar los resultados previos en términos de los procesos de contorno en lugar de los procesos de altura. La función de contorno de un árbol se discutió en la sección 1.1. A diferencia del proceso de altura, es conveniente indexar la función de contorno con un parámetro real.

Nuestro primer cometido será dar el resultado análogo al teorema 1.3.2, para lo cual consideraremos una sucesión  $\theta_1, \theta_2, \dots$  de  $\text{GWT}(\mu)$  independientes y denotaremos por  $C = \{C_t : t \geq 0\}$  al proceso que se obtiene al concatenar las funciones de contorno de los árboles  $\theta_i$ . Para que lo anterior tenga sentido, debemos definir de manera precisa el sentido en el qué concatenamos. En la sección 1.1, ver la figura 1.1.1, definimos la función de altura de un árbol  $\theta$  en el intervalo de tiempo  $[0, \zeta(\theta)]$ , donde  $\zeta(\theta) = 2(\#\theta - 1)$ . Esta definición conlleva a que la función de contorno de un árbol que solo contiene la raíz es trivial. Por este problema haremos una convención, artificial, de definir a la función de contorno  $C_\theta$  en el intervalo de tiempo  $[0, \xi(\theta)]$ , donde  $\xi(\theta) = 2\#\theta - 1 = \zeta(\theta) + 1$ , tomando  $C_t = 0$  para  $t \in [\zeta(\theta), \xi(\theta)]$ . Entonces obtenemos  $C$  al concatenar las funciones  $\{C_{\theta_i}(t) : t \in [0, \xi(\theta_i)]\}$ ,  $i \in \mathbb{N}$ .

Ahora, para cada  $n \in \mathbb{N}_0$  definimos

$$J_n = 2n - H_n + I_n.$$

Observemos que  $H_{n+1} - H_n \leq 1$  con la igualdad dándose si el vértice que se visita en el paso  $n + 1$  es descendiente del  $n$ -ésimo vértice, mientras que  $I_n - I_{n+1} \leq 1$  se da debido a que el ínfimo de la caminata aleatoria decrece a lo más en uno. En vista de que  $I_n = I_{n+1} + 1$  precisamente cuando el vértice visitado en el paso  $n + 1$  pertenece a un árbol distintos al que pertenece en  $n$ -ésimo vértice, deducimos que  $H_{n+1} - H_n + I_n - I_{n+1} < 2$ . Por otra parte, notando que  $-I_n$  nos indica el índice del árbol en el cual se encuentra el  $n$ -ésimo vértice, obtenemos que  $H_n - I_n \leq n$ . De la primera desigualdad deducimos que  $J_n$  es estrictamente creciente y de la segunda que  $J_n \geq n$ .

Recordemos que el proceso de altura al tiempo  $n$  corresponde a la generación del individuo visitado en el  $n$ -ésimo paso, considerando que se visitan los individuos en el orden lexicográfico dentro de cada árbol. Entonces, tomando en cuenta la construcción realizada de la función de contorno en la sección 1.1, en particular las expresiones (1.1.1) y (1.1.2), vemos que  $[J_n, J_{n+1}]$  es el intervalo de tiempo en el cual

la función de contorno pasa del individuo  $n$  al individuo  $n + 1$ . De hecho, para  $t \in [J_n, J_{n+1}]$  tenemos que

$$C_t = \begin{cases} H_n - (t - J_n) & \text{si } t \in [J_n, J_{n+1} - 1], \\ \left( (H_{n+1} - (J_{n+1} - t)) \right)_+ & \text{si } t \in [J_{n+1} - 1, J_{n+1}]. \end{cases}$$

Por lo tanto, para  $t \in [J_n, J_{n+1} - 1]$  tenemos

$$\begin{aligned} |C_t - H_n| &= |t - J_n| \leq J_{n+1} - J_n - 1 \leq 1 + |H_{n+1} - H_n| + (I_{n+1} - I_n) \\ &\leq |H_{n+1} - H_n| + 1, \end{aligned}$$

mientras que para  $t \in [J_{n+1} - 1, J_{n+1}]$ ,

$$\begin{aligned} |C_t - H_n| &= \left| \left( (H_{n+1} - (J_{n+1} - t)) \right)_+ - (H_n)_+ \right| \leq |H_{n+1} - H_n| + (J_{n+1} - t) \\ &\leq |H_{n+1} - H_n| + 1. \end{aligned}$$

En consecuencia,

$$\sup_{t \in [J_n, J_{n+1}]} |C_t - H_n| \leq |H_{n+1} - H_n| + 1.$$

Por último, definamos la función aleatoria  $\varphi : \mathbb{R}_+ \rightarrow \mathbb{N}_0$  poniendo  $\varphi(t) = n$  si y sólo si  $t \in [J_n, J_{n+1})$ . Por las cotas derivadas hasta ahora tendremos, para  $m \in \mathbb{N}$ ,

$$\sup_{t \in [0, m]} |C_t - H_{\varphi(t)}| \leq \sup_{t \in [0, J_m]} |C_t - H_{\varphi(t)}| \leq 1 + \sup_{n \leq m} |H_{n+1} - H_n|.$$

De manera similar, usando la definición de  $J_n$  obtenemos, para  $t \in [J_n, J_{n+1})$ , la desigualdad  $2n - H_n + I_n \leq t \leq 2n + 2 - H_{n+1} + I_{n+1}$  que a su vez nos permite obtener las cotas

$$-2 - H_n + I_n \leq -2 + H_{n+1} - I_{n+1} \leq 2n - t \leq H_n - I_n \leq 2 + H_n - I_n,$$

mediante las cuales deducimos

$$\left| \varphi(t) - \frac{t}{2} \right| \leq 1 + \frac{1}{2}H_n + \frac{1}{2}|I_n|,$$

y por ende

$$\sup_{t \in [0, m]} \left| \varphi(t) - \frac{t}{2} \right| \leq \sup_{t \in [0, J_m]} \left| \varphi(t) - \frac{t}{2} \right| \leq \frac{1}{2} \sup_{n \leq m} H_n + \frac{1}{2}|I_m| + 1.$$

### 1. Teorema

*Se cumple la convergencia*

$$\left\{ \frac{1}{\sqrt{p}} C_{2pt} : t \geq 0 \right\} \xrightarrow[p \rightarrow \infty]{} \frac{2}{\sigma} |\beta|,$$

*con  $\beta$  un movimiento browniano estándar.*

Demostración: Para cada  $p \in \mathbb{N}$  sea  $\varphi_p(t) = \varphi(pt)/p$ . Dado  $m \in \mathbb{N}$  observamos que

$$(1) \quad \sup_{t \leq m} \left| \frac{1}{\sqrt{p}} C_{2pt} - \frac{1}{\sqrt{p}} H_{p\varphi_p(2t)} \right| \leq \frac{1}{\sqrt{p}} + \frac{1}{\sqrt{p}} \sup_{t \leq 2m} |H_{\lfloor pt \rfloor + 1} - H_{\lfloor pt \rfloor}| \xrightarrow[p \rightarrow \infty]{} 0,$$

donde la convergencia es en probabilidad debido al teorema 1.3.2.

Por su parte, el teorema de Donsker implica, para cada  $m \in \mathbb{N}$ ,

$$\frac{1}{\sqrt{p}} I_{mp} \xrightarrow[p \rightarrow \infty]{} \sigma B_m.$$

Así, obtenemos que

$$(2) \quad \sup_{t \leq m} |\varphi_p(2t) - t| \leq \frac{1}{p} \sup_{k \leq 2mp} H_k + \frac{1}{p} |I_{2mp}| + \frac{2}{p} \xrightarrow[p \rightarrow \infty]{} 0,$$

con la convergencia siendo en probabilidad. El resultado se sigue entonces del teorema 1.3.2 junto a las convergencias (1) y (2). ■

También el corolario 1.1 y el teorema 2.1 se pueden reescribir en términos del proceso de contorno. En efecto, hay que reemplazar  $H_{[pt]}^{(p)}$  por  $C_{[2pt]}^{(p)}$  y  $H_{[pt]}^{(p)}$  por  $C_{[2pt]}^{(p)}$  en el corolario y el teorema de forma respectiva. En el caso del análogo del teorema 2.1, una demostración alternativa que únicamente usa teoría de excursiones se presenta en el artículo «Itô's Excursion Theory and Random Trees» de Le Gall (2010).

Los resultados expuestos hasta en los capítulos 1 y 4 nos indican la convergencia de los procesos de altura de Galton–Watson a excursiones brownianas. Empero, esto no nos da la convergencia de los árboles como elementos aleatorios. Para estudiarlo se requiere definir de manera formal qué son los árboles aleatorios límite y en qué sentido es la convergencia. Ese será el contenido del capítulo 5.

## 4.A. Demostración del lema 2.2

Para cerrar el capítulo, presentamos la demostración del lema 2.2, que recordamos nos dice que una caminata aleatoria que comienza en 0, condicionada a tocar por primera vez el valor  $-1$  en el tiempo  $p$ , converge a una excursión browniana normalizada. Conviene destacar que si se considera el enfoque de Le Gall (2010), no es necesario usar este lema. No obstante, como en estas notas seguimos a Le Gall (2005), presentaremos la demostración, la cual seguirá de cerca la presentación del artículo de Kaigh (1976).

En sí, demostraremos que la colección de medidas  $Q_p$  dadas por las distribuciones de  $\{S_{[pt]}/\sigma\sqrt{p} : t \in [0, 1]\}$  bajo  $\mathbb{P}(\cdot | T_0 = p)$ , con  $T_0 = \inf\{n \in \mathbb{N} : S_n = 0\}$ , converge débilmente a la ley de una excursión browniana normalizada, lo cual es ligeramente distinto al enunciado del lema 2.2, pero equivalente debido al reescalamiento considerado en espacio y tiempo.

La demostración se dividirá en dos partes, primero veremos que se da la convergencia finito-dimensional y posteriormente comprobaremos que la sucesión de medidas  $Q_p$  es tensa. Previo a ello, estableceremos algunas convenciones notacionales y algunos resultados auxiliares que consideraremos en lo que resta de la sección.

Pongamos  $r_p = \mathbb{P}(T_0 = p)$  y  $f_p = \mathbb{P}(T_0 = p)$ . Es sabido, ver la demostración del teorema 8 en el artículo de Kesten (1963), que se cumplen tanto

$$(1) \quad \lim_{p \rightarrow \infty} \sqrt{p} r_p = \sqrt{\frac{2}{\pi}} \sigma$$

como

$$(2) \quad \lim_{p \rightarrow \infty} \sqrt{p^3} f_p = \frac{\sigma}{\sqrt{2\pi}}.$$

Otros dos resultados que serán de utilidad en la prueba son

$$(3) \quad \lim_{p \rightarrow \infty} \mathbb{P}(S_p \leq x\sqrt{p} \mid T_0 > p) = \int_{-\infty}^x \frac{|y|}{2\sigma^2} e^{-y^2/2\sigma^2} dy$$

y, una versión local,

$$(4) \quad \lim_{p \rightarrow \infty} \sup_{x \in \mathbb{Z}} \left| p \mathbb{P}(T_{\{x\}} = p) - \frac{|x|}{\sigma\sqrt{p}} \varphi\left(\frac{x}{\sigma\sqrt{p}}\right) \right| = 0,$$

donde  $T_{\{x\}} = \inf\{n \in \mathbb{N} : S_n = x\}$  y  $\varphi$  es la función de densidad de la distribución normal estándar. Las referencias de estos resultados se encuentran dentro del artículo de Kaigh (1976).

En la demostración será útil usar la caracterización de la excursión browniana normalizada  $e$  dada en la sección 2.9 del libro *Diffusion Processes and Their Sample Paths* de Ito y McKean (1996). Esta caracterización nos dice que  $e$  es un proceso de Markov en  $[0, 1]$  con densidad de transición no estacionaria  $q$  dada por

$$\begin{aligned} q(0, 0; t, y) dy &= \mathbb{P}(e_t \in dy) \\ &= \frac{1}{\sqrt{2\pi t^3(1-t)^3}} y^2 \exp\left\{-\frac{y^2}{2t(1-t)}\right\} dy \end{aligned}$$

para  $t \in (0, 1)$  y  $y \in \mathbb{R}$ , mientras que

$$\begin{aligned} q(t_1, y_1; t_2, y_2) dy_2 &= \mathbb{P}(e_{t_2} \in dy_2 \mid e_{t_1} = y_1) \\ &= \left(\frac{1-t_1}{1-t_2}\right)^{3/2} \frac{y_2}{y_1} \frac{\varphi(y_2/\sqrt{1-t_2})}{\varphi(y_1/\sqrt{1-t_1})} \frac{1}{\sqrt{t_2-t_1}} \\ &\quad \times \left[ \varphi\left(\frac{y_1-y_2}{\sqrt{t_2-t_1}}\right) - \varphi\left(\frac{y_1+y_2}{\sqrt{t_2-t_1}}\right) \right] \end{aligned}$$

para  $0 < t_1 < t_2 < 1$  y  $y_1, y_2 > 0$ .

**Demostración de la convergencia finito-dimensional** Comenzamos observando que se tiene

$$\begin{aligned} \int_{-\infty}^x \frac{1}{r_p} \mathbb{P}\left(\frac{S_{\lfloor pt \rfloor}}{\sigma\sqrt{p}} \in dy, T_0 > \lfloor pt \rfloor\right) &= \frac{r_{\lfloor pt \rfloor}}{r_p} \mathbb{P}(S_{\lfloor pt \rfloor} \leq \sigma x\sqrt{p} \mid T_0 > \lfloor pt \rfloor) \\ &\xrightarrow{p \rightarrow \infty} \frac{1}{\sqrt{t}} \int_{-\infty}^{\sigma x/\sqrt{t}} \frac{|y|}{2\sigma^2} e^{-y^2/2\sigma^2} dy \\ &= \int_{-\infty}^x \frac{|z|}{2t^{3/2}} e^{-z^2/2t} dz, \end{aligned}$$

donde la convergencia se da usando (1) y (3) y la última igualdad por el cambio de variable  $z = \sqrt{t}y/\sigma$ . Es decir que se tiene

$$(5) \quad \lim_{p \rightarrow \infty} \int_{-\infty}^x \frac{1}{r_p} \mathbb{P}\left(\frac{S_{\lfloor pt \rfloor}}{\sigma\sqrt{p}} \in dy, T_0 > \lfloor pt \rfloor\right) = \int_{-\infty}^x \frac{|y|}{2t^{3/2}} e^{-y^2/2t} dy.$$

Por otro lado, Kaigh (1976) menciona el siguiente resultado de Belkin (1972):

$$(6) \quad \begin{aligned} & \lim_{p \rightarrow \infty} \int_{-\infty}^x \mathbb{P}_{\sigma\sqrt{p}y_1} \left( \frac{S_{\lfloor pt \rfloor}}{\sigma\sqrt{p}} \in dy_2, T > \lfloor pt \rfloor \right) \\ &= \int_{-\infty}^x \frac{1}{\sqrt{t}} \left[ \varphi \left( \frac{y_1 - y_2}{\sqrt{t}} \right) - \varphi \left( \frac{y_1 + y_2}{\sqrt{t}} \right) \right] \mathbb{1}_{\{y_1 y_2 > 0\}} dy_2 \end{aligned}$$

de manera uniforme en compactos como función de  $y_1$ . Dicho lo anterior, notemos ahora que

$$\frac{r_p}{f_p} \mathbb{P}^{\sigma\sqrt{p}y}(T_0 = p - \lfloor pt \rfloor) = \frac{r_p}{pf_p} \frac{p}{p - \lfloor pt \rfloor} (p - \lfloor pt \rfloor) \mathbb{P}(T_{\{-\sigma\sqrt{p}y\}} = p - \lfloor pt \rfloor).$$

Por (1) y (2) deducimos

$$\lim_{p \rightarrow \infty} \frac{r_p}{pf_p} = 2,$$

mientras que es elemental que

$$\lim_{p \rightarrow \infty} \frac{p}{p - \lfloor pt \rfloor} = \frac{1}{1 - t}.$$

Por su parte, poniendo  $v(p) = p - \lfloor pt \rfloor$ , tendremos

$$\begin{aligned} v(p) \mathbb{P}(T_{\{-\sigma\sqrt{p}y\}} = v(p)) &= \left( v(p) \mathbb{P}(T_{\{-\sigma\sqrt{p}y\}} = v(p)) - \frac{\sigma\sqrt{p}|y|}{\sigma\sqrt{v(p)}} \varphi \left( \frac{\sigma\sqrt{p}|y|}{\sigma\sqrt{v(p)}} \right) \right) \\ &\quad + \sqrt{\frac{p}{v(p)}} |y| \varphi \left( \sqrt{\frac{p}{v(p)}} |y| \right). \end{aligned}$$

En vista de (4), el primer término converge a cero, mientras que es claro que

$$\lim_{p \rightarrow \infty} \sqrt{\frac{p}{v(p)}} |y| \varphi \left( \sqrt{\frac{p}{v(p)}} |y| \right) = \frac{|y|}{\sqrt{1 - t}} \varphi \left( \frac{|y|}{\sqrt{1 - t}} \right).$$

Así, hemos probado que

$$(7) \quad \lim_{p \rightarrow \infty} \frac{r_p}{f_p} \mathbb{P}^{\sigma\sqrt{p}y}(T_0 = p - \lfloor pt \rfloor) = \frac{2|y|}{(1 - t)^{3/2}} \varphi \left( \frac{|y|}{\sqrt{1 - t}} \right)$$

de manera uniforme en  $y$ .

La convergencia finito-dimensional se sigue entonces mediante las expresiones (5), (6) y (7) junto a un argumento recursivo análogo al realizado por Belkin (1972, pág. 54).

**Convergencia funcional (tensión)** Buscaremos aplicar el corlario de la página 142 en el libro de Billingsley (1999). Debido a que ya tenemos la convergencia finito-dimensional basta verificar las condiciones:

(i) Para cada  $\eta > 0$  existen  $a > 0$  y  $n_0$  tales que, para  $p \geq n_0$

$$Q_p(|\omega(0)| \geq a) \leq \eta.$$

(ii) Para cualesquiera  $\varepsilon, \eta > 0$  existen  $\delta \in (0, 1)$  y  $n_0$  bajo los cuales

$$Q_p(w(\omega, \delta) \geq \varepsilon) \leq \eta$$

siempre que  $p \geq n_0$ , donde

$$w_\omega(\delta) = w(\omega, \delta) = \sup_{|s-t| \leq \delta} |\omega(s) - \omega(t)|$$

es el módulo de continuidad de  $\omega$ .

Puesto que  $Q_p(\{0\}) = 1$ , la condición (i) se satisface trivialmente. Resulta suficiente probar que se satisface la condición (ii), lo cual es equivalente a demostrar

$$(8) \quad \lim_{\delta \rightarrow 0} \limsup_{p \rightarrow \infty} Q_p(w(\omega, \delta) \geq 2\varepsilon) = 0$$

para cualquier  $\varepsilon > 0$ . Para lograr nuestro cometido requeriremos un par de resultados auxiliares.

### 1. Lema

Dado  $\varepsilon > 0$ , para cada  $\eta \in (0, 1)$  existe un entero positivo  $n_0 = n_0(\eta, \varepsilon)$  para el cual se tiene

$$\begin{aligned} \mathbb{P}^{\sqrt{p}x}(T_0 = p - k) &\leq 4\sqrt{2\pi} \frac{\sigma}{\varepsilon} e^{\varepsilon^2/2\sigma^2(1-\eta)} \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor) \\ &= c(\eta, \varepsilon) \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor) \end{aligned}$$

para cualquier elección de  $x$  y  $k \leq \lfloor p\eta \rfloor$  siempre que  $p > n_0$ .

Demostración: La expresión (4) implica la existencia de  $N_1 \in \mathbb{N}$  para el cual, siempre que  $p > N_1$ , se tiene

$$\sup_{x \in \mathbb{R}} \left| p \mathbb{P}^{\sqrt{p}x}(T_0 = p) - \frac{|x|}{\sigma} \varphi\left(\frac{x}{\sigma}\right) \right| < 1.$$

Escojamos  $n_1$  de tal forma que  $p - \lfloor p\eta \rfloor > N_1$  cuando  $p > n_1$ . Entonces, para  $k \leq \lfloor p\eta \rfloor$ , con  $p > n_1$ ,

$$\begin{aligned} \sup_{x \in \mathbb{R}} (p - k) \mathbb{P}^{\sqrt{p}x}(T_0 = p - k) &\leq 1 + \sup_{x \in \mathbb{R}} \frac{\sqrt{n}|x|}{\sigma\sqrt{n-k}} \varphi\left(\frac{\sqrt{n}x}{\sigma\sqrt{n-k}}\right) \\ &\leq 1 + \sup_{y \in \mathbb{R}} |y| \varphi(y) = 1 + \varphi(1) \leq 2. \end{aligned}$$

De la prueba de (7) sabemos que

$$\lim_{p \rightarrow \infty} (p - \lfloor p\eta \rfloor) \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor) = \frac{\varepsilon}{\sigma\sqrt{1-\eta}} \varphi\left(\frac{\varepsilon}{\sigma\sqrt{1-\eta}}\right),$$

de modo que existe  $n_2 \in \mathbb{N}$  tal que

$$(p - \lfloor p\eta \rfloor) \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor) \geq \frac{\varepsilon}{2\sigma\sqrt{1-\eta}} \varphi\left(\frac{\varepsilon}{\sigma\sqrt{1-\eta}}\right)$$

para  $p > n_2$ . Consecuentemente, si  $p > n_1 \vee n_2$  y  $k \leq \lfloor p\eta \rfloor$ ,

$$\begin{aligned} \frac{\sup_x \mathbb{P}^{\sqrt{p}x}(T_0 = p - k)}{\mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor)} &\leq \frac{p - \lfloor p\eta \rfloor}{p - k} \frac{2}{(\varepsilon/2\sigma)(1-\eta)^{-1/2} \varphi(\varepsilon/\sigma\sqrt{1-\delta})} \\ &\leq \frac{4\sigma}{\varepsilon} \sqrt{1-\eta} \varphi\left(\frac{\varepsilon}{\sigma\sqrt{1-\delta}}\right) \leq c(\eta, \varepsilon), \end{aligned}$$

con lo que obtenemos el resultado deseado. ■

El siguiente resultado que se presenta es una consecuencia de la ecuación (2.10) y la discusión posterior en el artículo de Belkin (1972, pág. 49).

## 2. Lema

Dado  $\varepsilon > 0$ , para cada  $\eta \in (0, 1)$  existen  $\lambda(\eta, \varepsilon) > 0$  y  $m_0 = m_0(\eta, \varepsilon)$  tales que

$$\mathbb{P}^{\sqrt{p}x}(S_{\lfloor p\eta \rfloor - k} \geq \sqrt{p}x, T_0 > \lfloor p\eta \rfloor - k) \geq \frac{1}{2}\lambda(\eta, \varepsilon)$$

para cada  $x \geq \varepsilon$ , siempre que  $\lfloor n\eta \rfloor - k \geq m_0$ . En adición,  $\lim_{\eta \rightarrow 0} \lambda(\eta, \varepsilon) > 0$ .

Es momento de proceder con la demostración de (8), para lo cual definiremos el tiempo aleatorio

$$\tau = \tau(\delta, \varepsilon) = \inf\left\{t \in [0, 1] : \sup_{s \in [t, t+\delta]} |\omega(s) - \omega(t)| \geq 2\varepsilon\right\},$$

donde  $\inf \emptyset = \infty$ . Para reducir la notación un poco, para  $A \subset [0, 1]$  pondremos

$$w(\omega, A) = w_\omega(A) = \sup_{s, t \in A} |\omega(s) - \omega(t)|.$$

Entonces, para  $\eta \in (0, 1/2)$  tendremos que

$$Q_p(w(\omega, \delta) \geq 2\varepsilon) = Q_p(\tau \leq \eta) + Q_p(\eta < \tau < 1 - \eta) + Q_p(1 - \eta \leq \tau \leq 1)$$

y por consiguiente bastará demostrar que cada uno de los términos en

$$\limsup_{\delta \rightarrow 0} \limsup_{p \rightarrow \infty} Q_p(\tau \leq \eta),$$

$$\limsup_{\delta \rightarrow 0} \limsup_{p \rightarrow \infty} Q_p(\eta < \tau < 1 - \eta),$$

y

$$\limsup_{\delta \rightarrow 0} \limsup_{p \rightarrow \infty} Q_p(1 - \eta \leq \tau \leq 1),$$

convergen a cero conforme  $\eta \rightarrow 0$ .

Comenzamos considerando el término que corresponde a  $Q_p(\eta < \tau < 1 - \eta)$ , en este caso observamos que

$$\begin{aligned} Q_p(\eta < \tau < 1 - \delta) &= \sum_{x \neq 0} \frac{1}{f_p} \mathbb{P}(\eta < \tau < 1 - \eta, S_{p - \lfloor p\eta \rfloor} = \sqrt{p}x, T_0 = p) \\ &= \sum_{x \neq 0} \frac{1}{f_p} \mathbb{P}(\eta < \tau < 1 - \eta, S_{p - \lfloor p\eta \rfloor} = \sqrt{p}x, T_0 = p, T_0 > p - \lfloor \eta p \rfloor) \\ &= \sum_{x \neq 0} \frac{1}{f_p} \mathbb{P}(\eta < \tau < 1 - \eta, S_{p - \lfloor p\eta \rfloor} = \sqrt{p}x, T_0 > p - \lfloor p\eta \rfloor) \mathbb{P}^{\sqrt{p}x}(T_0 = \lfloor p\eta \rfloor), \end{aligned}$$

donde la última igualdad se da por la propiedad de Markov. Entonces, al usar el lema 1 (con  $1 - \eta$  en lugar de  $\eta$ ), deducimos que para  $n$  suficientemente grande se cumplirá

$$\sup_x \mathbb{P}^{\sqrt{p}x}(T_0 = \lfloor p\eta \rfloor) \leq c(1 - \eta, \varepsilon) \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = \lfloor p\eta \rfloor).$$

Por consiguiente deducimos

$$\begin{aligned}
Q_p(\eta < \tau < 1 - \eta) &\leq \frac{c(1 - \eta, \varepsilon)}{f_p} \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = \lfloor p\eta \rfloor) \mathbb{P}(\delta < \tau < 1 - \eta, T_0 > p - \lfloor p\eta \rfloor) \\
&\leq \frac{c(1 - \eta, \varepsilon)}{f_p} \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = \lfloor p\eta \rfloor) \mathbb{P}(\eta < \tau < 1 - \eta, T_0 > \lfloor p\eta \rfloor) \\
&\leq \frac{c(1 - \eta, \varepsilon)}{f_p} \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = \lfloor p\eta \rfloor) \mathbb{P}\left(\sup_{\eta \leq s \leq 1 - \eta} w_{p^{-1/2}S_{\lfloor p \cdot \rfloor}}([s, s + \rho]) \geq 2\varepsilon, T_0 > \lfloor p\eta \rfloor\right) \\
&\leq \frac{c(1 - \eta, \varepsilon)}{f_p} \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = \lfloor p\eta \rfloor) \mathbb{P}(T_0 > \lfloor p\eta \rfloor) \mathbb{P}\left(w(p^{-1/2}S_{\lfloor p \cdot \rfloor}, \rho) \geq 2\varepsilon\right)
\end{aligned}$$

con la segunda igualdad dándose por tomar  $\eta < 1/2$ . Procediendo como se hizo para obtener (7) es sencillo verificar que

$$\lim_{p \rightarrow \infty} \frac{r_{\lfloor p\eta \rfloor}}{f_p} \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = \lfloor p\eta \rfloor) = \frac{2\varepsilon}{\sigma\eta^2} \varphi\left(\frac{\varepsilon}{\sigma\eta}\right) < \infty.$$

Por otra parte, el teorema de Donsker implica que  $\mathbb{P}\left(w(p^{-1/2}S_{\lfloor p \cdot \rfloor}, \rho) \geq 2\varepsilon\right) \rightarrow 0$  conforme  $p \rightarrow \infty$ , con lo cual obtenemos

$$\limsup_{p \rightarrow \infty} Q_p(\eta < \tau < 1 - \eta) = 0 \quad \text{para cualquier } \eta \in (0, 1/2).$$

Ahora trabajaremos sobre el término  $Q_p(\tau \leq \eta)$ . Por definición se cumple

$$Q_p(\tau \leq \eta) \leq \mathbb{P}\left(\sup_{t \leq \eta} |S_{\lfloor pt \rfloor}| \geq \sqrt{p}\varepsilon \mid T_0 = p\right),$$

y en consecuencia basta demostrar que

$$\limsup_{p \rightarrow \infty} \frac{1}{f_p} \mathbb{P}\left(\sup_{t \leq \eta} |S_{\lfloor pt \rfloor}| \geq \sqrt{p}\varepsilon, T_0 = p\right) \rightarrow 0 \quad \text{conforme } \eta \rightarrow 0.$$

Para lograr este cometido definamos  $\tau_p = \min\{j \in [\lfloor p\eta \rfloor] : S_j \geq \sqrt{p}\varepsilon\}$ . Dado  $m < \lfloor p\eta \rfloor$ , una aplicación de la propiedad de Markov y lema 1 nos dan la primera desigualdad en

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t \leq \eta} S_{\lfloor pt \rfloor} \geq \sqrt{p}\varepsilon, \tau_p \geq \lfloor p\eta \rfloor - m, T_0 = p\right) \\
&= \sum_{k=p\eta-m}^{p\eta} \sum_{x \geq \varepsilon} \mathbb{P}(\tau_p = k, S_k = \sqrt{p}x, T_0 = p, T_0 > k) \\
&\leq c(\eta, \varepsilon) \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor) \sum_{k=p\eta-m}^{p\eta} \sum_{x \geq \varepsilon} \mathbb{P}(\tau_p = k, S_k = \sqrt{p}x, T_0 > k).
\end{aligned}$$

Notemos que para  $k \in [\lfloor p\eta \rfloor - m, \lfloor p\eta \rfloor]$ ,  $m \geq \lfloor p\eta \rfloor - k$  y entonces, para todo  $x \geq \varepsilon$ ,

$$\begin{aligned}
\mathbb{P}\left(\inf_{j \leq m} S_j \geq -\frac{\sqrt{p}\varepsilon}{2}\right) &\leq \mathbb{P}\left(\inf_{j \leq \lfloor p\eta \rfloor - k} S_j \geq -\frac{\sqrt{p}\varepsilon}{2}\right) \\
&= \mathbb{P}^{\sqrt{p}x}\left(\inf_{j \leq \lfloor p\eta \rfloor - k} \frac{S_j}{\sqrt{p}} \geq x - \frac{\varepsilon}{2}\right) \\
&\leq \mathbb{P}^{\sqrt{p}x}\left(\frac{S_{\lfloor p\eta \rfloor - k}}{\sqrt{p}} \geq x - \frac{\varepsilon}{2}, T_0 > \lfloor p\eta \rfloor - k\right)
\end{aligned}$$

Así deducimos que para  $n$  suficientemente grande,

$$\begin{aligned}
& \mathbb{P}\left(\sup_{t \leq \eta} S_{\lfloor pt \rfloor} \geq \sqrt{p}\varepsilon, \tau_p \geq \lfloor p\eta \rfloor - m, T_0 = p\right) \\
& \leq c(\eta, \varepsilon) \frac{\mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor)}{\mathbb{P}(\inf_{j \leq m} S_j \geq -\sqrt{p}\varepsilon/2)} \sum_{k=p\eta-m}^{p\eta} \sum_{x \geq \varepsilon} \mathbb{P}(\tau_p = k, S_k = \sqrt{p}x, T_0 > k) \\
& \quad \times \mathbb{P}^{\sqrt{p}x}\left(\frac{S_{\lfloor p\eta \rfloor - k}}{\sqrt{p}} \geq x - \frac{\varepsilon}{2}, T_0 > \lfloor p\eta \rfloor - k\right) \\
& = c(\eta, \varepsilon) \frac{\mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor)}{\mathbb{P}(\inf_{j \leq m} S_j \geq -\sqrt{p}\varepsilon/2)} \mathbb{P}\left(S_{\lfloor p\eta \rfloor} \geq \frac{\sqrt{p}\varepsilon}{2}, \tau_p \geq \lfloor p\eta \rfloor - m, T_0 > \lfloor p\eta \rfloor\right),
\end{aligned}$$

donde hemos usado la propiedad de Markov en la última igualdad. Ahora obtendremos una desigualdad similar cuando  $\tau_p < \lfloor p\eta \rfloor - m$  para lo cual observamos que

$$\begin{aligned}
& \mathbb{P}\left(\sup_{t \leq \eta} S_{\lfloor pt \rfloor} \geq \sqrt{p}\varepsilon, \tau_p < \lfloor p\eta \rfloor - m, T_0 = p\right) \\
& = \sum_{k=1}^{p\eta-m-1} \sum_{x \geq \varepsilon} \mathbb{P}(\tau_p = k, S_k = \sqrt{p}x, T_0 > k) \mathbb{P}^{\sqrt{p}x}(T_0 = p - k),
\end{aligned}$$

de modo que los lemas 1 y 2 nos permiten obtener, para  $n$  suficientemente grande, la primera desigualdad en

$$\begin{aligned}
& \mathbb{P}\left(\sup_{t \leq \eta} S_{\lfloor pt \rfloor} \geq \sqrt{p}\varepsilon, \tau_p < \lfloor p\eta \rfloor - m, T_0 = p\right) \\
& \leq \frac{2c(\eta, \varepsilon)}{\lambda(\eta, \varepsilon)} \sum_{k=1}^{p\eta-m-1} \sum_{x \geq \varepsilon} \mathbb{P}(\tau_p = k, S_k = \sqrt{p}x, T_0 > k) \\
& \quad \times \mathbb{P}^{\sqrt{p}x}(S_{\lfloor p\eta \rfloor - k} \geq \sqrt{p}x, T_0 > \sqrt{p}\eta - k) \\
& \quad \times \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor) \\
& \leq \frac{2c(\eta, \varepsilon)}{\lambda(\eta, \varepsilon)} \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor) \mathbb{P}\left(S_{\lfloor p\eta \rfloor} > \sqrt{p}\varepsilon, \tau_p < \lfloor p\eta \rfloor - m, T_0 > \lfloor p\eta \rfloor\right).
\end{aligned}$$

Esto nos permite obtener la cota

$$\begin{aligned}
& \mathbb{P}\left(\sup_{t \leq \eta} S_{\lfloor pt \rfloor} \geq \sqrt{p}\varepsilon, T_0 = p\right) \\
& \leq c(\eta, \varepsilon) \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor) \mathbb{P}\left(\frac{S_{\lfloor p\eta \rfloor}}{\sqrt{p}} \geq \frac{\varepsilon}{2}, T_0 > \lfloor p\eta \rfloor\right) \\
& \quad \times \left[ \frac{2}{\lambda(\eta, \varepsilon)} + \frac{1}{\mathbb{P}(\inf_{j \leq m} S_j \geq -\sqrt{p}\varepsilon/2)} \right].
\end{aligned}$$

En vista de (4) obtenemos

$$\lim_{p \rightarrow \infty} p \mathbb{P}^{\sqrt{p}\varepsilon}(T_0 = p - \lfloor p\eta \rfloor) = \frac{\varepsilon}{\sigma(1-\eta)^{3/2}} \varphi\left(\frac{\varepsilon}{\sigma(1-\eta)^{1/2}}\right),$$

mientras que de (1) y (3) obtenemos

$$\begin{aligned}
\lim_{p \rightarrow \infty} \sqrt{p} \mathbb{P}\left(\frac{S_{\lfloor p\eta \rfloor}}{\sqrt{p}} \geq \frac{\varepsilon}{2}, T_0 > \lfloor p\eta \rfloor\right) &= \sqrt{\frac{2\sigma^2}{\pi\eta}} \int_{\varepsilon/2\sqrt{\eta}}^{\infty} \frac{|y|}{2\sigma^2} e^{-y^2/2\sigma^2} dy \\
&= \frac{1}{\sqrt{2\pi\eta/\sigma^2}} e^{-\varepsilon^2/8\sigma^2\eta}.
\end{aligned}$$

En tanto,  $\lim_{p \rightarrow \infty} \mathbb{P}(\inf_{j \leq m} S_j \geq -\sqrt{p}\varepsilon/2) = 1$  debido a que  $m$  es fijo. De estas relaciones y (2) obtenemos

$$\begin{aligned} & \limsup_{p \rightarrow \infty} \mathbb{P}(\sup_{t \leq \eta} S_{\lfloor pt \rfloor} \geq \sqrt{p}\varepsilon \mid T_0 = p) \\ & \leq c(\eta, \varepsilon) \left[ \frac{2}{\lambda(\eta, \varepsilon)} + 1 \right] \frac{\varepsilon}{\sqrt{\eta(1-\eta)^3 \sigma^2}} e^{-\varepsilon^2/8\sigma^2\eta} \varphi\left(\frac{\varepsilon}{\sigma\sqrt{1-\eta}}\right). \end{aligned}$$

Puesto que para  $\varepsilon > 0$  dado,  $\lim_{\eta \rightarrow 0} c(\eta, \varepsilon) < \infty$  y  $\lim_{p \rightarrow \infty} \lambda(\eta, \varepsilon) > 0$  concluimos que

$$\lim_{\eta \rightarrow 0} \limsup_{p \rightarrow \infty} \mathbb{P}(\sup_{t \leq \delta} S_{\lfloor pt \rfloor} \geq \sqrt{p}\varepsilon \mid T_0 = p) = 0.$$

Un argumento análogo, con  $-S$ , nos permite concluir

$$\lim_{\eta \rightarrow 0} \limsup_{p \rightarrow \infty} \mathbb{P}(\sup_{t \leq \eta} |S_{\lfloor pt \rfloor}| \geq \sqrt{p}\varepsilon \mid T_0 = p) = 0$$

y en consecuencia

$$\limsup_{\rho \rightarrow 0} \limsup_{p \rightarrow \infty} Q_p(\tau \leq \eta) \rightarrow 0$$

conforme  $\eta \rightarrow 0$ .

Para terminar la demostración basta considerar la caminata aleatoria dual  $\hat{S}_j^{(p)} = S_{p-j}$ , pues condicional en  $T_0 = p$ , ella tiene la misma distribución que  $\{S_j : 0 \leq j \leq p\}$  y obtenemos

$$\limsup_{\rho \rightarrow 0} \limsup_{p \rightarrow \infty} Q_p(1 - \eta \leq \tau \leq 1) \rightarrow 0$$

conforme  $\eta \rightarrow 0$  por lo demostrado previamente.

# Árboles reales y su codificación mediante excursiones brownianas

# 5

A lo largo de este capítulo nos dedicaremos a los árboles reales. Primeramente describiremos el formalismo de estos, para dar un significado preciso de la convergencia de árboles discretos reescalados a objetos continuos. Posteriormente veremos cómo codificar un árbol real mediante una función continua, de forma similar a la caracterización de los árboles discretos mediante su función de contorno. Por último veremos el árbol aleatorio continuo de Aldous (CRT por sus siglas en inglés), que puede ser definido como un árbol real codificado mediante una excursión browniana normalizada. Lo expuesto en el presente apartado se toma de la sección 2 del artículo de revisión de Le Gall (2005).

## 5.1. Árboles reales

Como bien se dijo en la introducción del capítulo, nuestro primer objetivo es introducir las nociones correspondientes a los árboles continuos. Comenzamos con una definición, en la cual incluiremos la propiedad de compacidad debido a que únicamente consideraremos árboles reales compactos.

### 1. Definición (Árbol métrico compacto)

Un espacio métrico compacto  $(\mathcal{T}, d)$  es un *árbol real* si para cualesquiera  $a, b \in \mathcal{T}$  se cumplen las propiedades:

- (i) Existe una única isometría  $f_{a,b}$  de  $[0, d(a, b)]$  en  $\mathcal{T}$  tal que  $f_{a,b}(0)$  y  $f_{a,b}(d(a, b)) = b$ .
- (ii) Si  $q_{a,b}$  es una función continua e inyectiva de  $[0, 1]$  a  $\mathcal{T}$  con  $q_{a,b}(0) = a$  y  $q_{a,b}(1) = b$ , entonces

$$q_{a,b}([0, 1]) = f_{a,b}([0, d(a, b)]).$$

Un *árbol real enraizado* es un árbol real  $(\mathcal{T}, d)$  con un vértice distinguido  $\rho = \rho(\mathcal{T})$ , el cual se conoce como *raíz*.

En lo que sigue, los árboles reales siempre tendrán una raíz. Incluso si no se menciona explícitamente.

Consideremos un árbol real  $(\mathcal{T}, d)$ . El rango del mapeo  $f_{a,b}$ , es decir el segmento de línea entre  $a$  y  $b$  en el árbol, se denota por  $\llbracket a, b \rrbracket$ . En particular,  $\llbracket \rho, a \rrbracket$  es el camino que va de la raíz a  $a$ , que interpretaremos como la línea ancestral del vértice  $a$ . Con esta noción podemos definir un orden parcial sobre el árbol, poniendo  $a \preceq b$  ( $a$  es ancestro de  $b$ ) si y sólo si  $a \in \llbracket \rho, b \rrbracket$ .

Por definición de árbol real, dados  $a, b \in \mathcal{T}$ , existe un único  $c \in \mathcal{T}$  tal que

$\llbracket \rho, a \rrbracket \cap \llbracket \rho, b \rrbracket = \llbracket \rho, c \rrbracket$ . Escribiremos  $c = a \wedge b$  y diremos que  $c$  es el *ancestro común más reciente* de  $a$  y  $b$ .

La multiplicidad de un vértice  $a \in \mathcal{T}$  es, por definición, el número de componentes conexas de  $\mathcal{T} \setminus \{a\}$ . Los vértices de  $\mathcal{T} \setminus \{\rho\}$  cuya multiplicidad sea 1 serán conocidos como *hojas*.

Puesto que nuestro objetivo principal es estudiar la convergencia de árboles reales aleatorios, es necesaria una noción de distancia entre dos árboles. Para ello usaremos la distancia de Gromov–Hausdorff. Una descripción detallada se puede consultar en el capítulo 7 del libro de Burago et al. (2001); aquí daremos una breve introducción.

Dado  $(\mathcal{X}, d_{\mathcal{X}})$  un espacio métrico, usamos la notación  $d_{\text{H}}(A, B)$  para la métrica de Hausdorff entre subconjuntos compactos  $A, B \subset \mathcal{X}$ , dada por

$$d_{\text{H}}(A, B) = \inf\{\varepsilon > 0 : A \subset B^{\varepsilon} \text{ y } B \subset A^{\varepsilon}\},$$

donde  $A^{\varepsilon} := \{x \in \mathcal{X} : d_{\mathcal{X}}(x, A) \leq \varepsilon\}$ .

Luego, si  $\mathcal{X}$  y  $\mathcal{Y}$  son dos espacios métricos compactos enraizados, con raíces  $\rho_{\mathcal{X}}$  y  $\rho_{\mathcal{Y}}$  de forma respectiva, entonces la distancia de Gromov–Hausdorff entre  $\mathcal{X}$  y  $\mathcal{Y}$  se denotará por  $d_{\text{GH}}(\mathcal{X}, \mathcal{Y})$  y se encontrará dada por

$$d_{\text{GH}}(\mathcal{X}, \mathcal{Y}) = \inf\{d_{\text{H}}(\varphi(\mathcal{X}), \varphi'(\mathcal{Y})) \vee d_{\mathcal{Z}}(\varphi(\rho_{\mathcal{X}}), \varphi'(\rho_{\mathcal{Y}}))\},$$

donde el ínfimo se toma sobre todas las elecciones de espacios métricos  $(\mathcal{Z}, d_{\mathcal{Z}})$  y todas las inserciones isométricas  $\varphi : \mathcal{X} \hookrightarrow \mathcal{Z}$  y  $\varphi' : \mathcal{Y} \hookrightarrow \mathcal{Z}$ . En la figura 1 se muestran los elementos descritos para definir la distancia de Gromov–Hausdorff.

Diremos que dos espacios métricos compactos y enraizados  $\mathcal{T}_1$  y  $\mathcal{T}_2$  son equivalentes si existe una isometría sobreyectiva de  $\mathcal{T}_1$  a  $\mathcal{T}_2$  que preserve la raíz. Dicho lo anterior, es claro que  $d_{\text{GH}}(\mathcal{X}, \mathcal{Y})$  únicamente depende de las clases de equivalencia de espacios métricos compactos enraizados. De ahora en adelante denotaremos por  $\mathbb{T}$  al conjunto de todos los árboles reales enraizados.

## 2. Teorema

*El espacio métrico  $(\mathbb{T}, d_{\text{GH}})$  es completo y separable.*

Una demostración de este teorema se puede consultar en el Teorema 1 del artículo de Evans et al. (2006), misma que reproduciremos por completud.

Antes de demostrar el teorema 2, haremos algunas observaciones. Es posible definir la distancia de Gromov–Hausdorff de manera distinta. Recordemos que si  $(\mathcal{X}, d_{\mathcal{X}})$  y  $(\mathcal{Y}, d_{\mathcal{Y}})$  son dos espacios métricos compactos, una relación de correspondencia entre  $\mathcal{X}$  y  $\mathcal{Y}$  es un subconjunto  $\mathcal{R}$  de  $\mathcal{X} \times \mathcal{Y}$  con la propiedad de que para cualquier  $x_1 \in \mathcal{X}$  existe  $y_1 \in \mathcal{Y}$  tal que  $(x_1, y_1) \in \mathcal{R}$  y de manera recíproca para cada  $y_2 \in \mathcal{Y}$ , existe  $x_2 \in \mathcal{X}$  tal que  $(x_2, y_2) \in \mathcal{R}$ . La *distorsión de la correspondencia*  $\mathcal{R}$  se define mediante

$$\text{dis}(\mathcal{R}) = \sup\{|d_{\mathcal{X}}(x_1, x_2) - d_{\mathcal{Y}}(y_1, y_2)| : (x_1, y_1), (x_2, y_2) \in \mathcal{R}\}.$$

Entonces, si  $\mathcal{X}$  y  $\mathcal{Y}$  son dos espacios métricos enraizados, con raíces  $\rho_{\mathcal{X}}$  y  $\rho_{\mathcal{Y}}$ , se cumple que

$$(1) \quad d_{\text{GH}}(\mathcal{X}, \mathcal{Y}) = \frac{1}{2} \inf_{\mathcal{R} \in \mathcal{C}(\mathcal{X}, \mathcal{Y}) : (\rho_{\mathcal{X}}, \rho_{\mathcal{Y}}) \in \mathcal{R}} \text{dis}(\mathcal{R}),$$

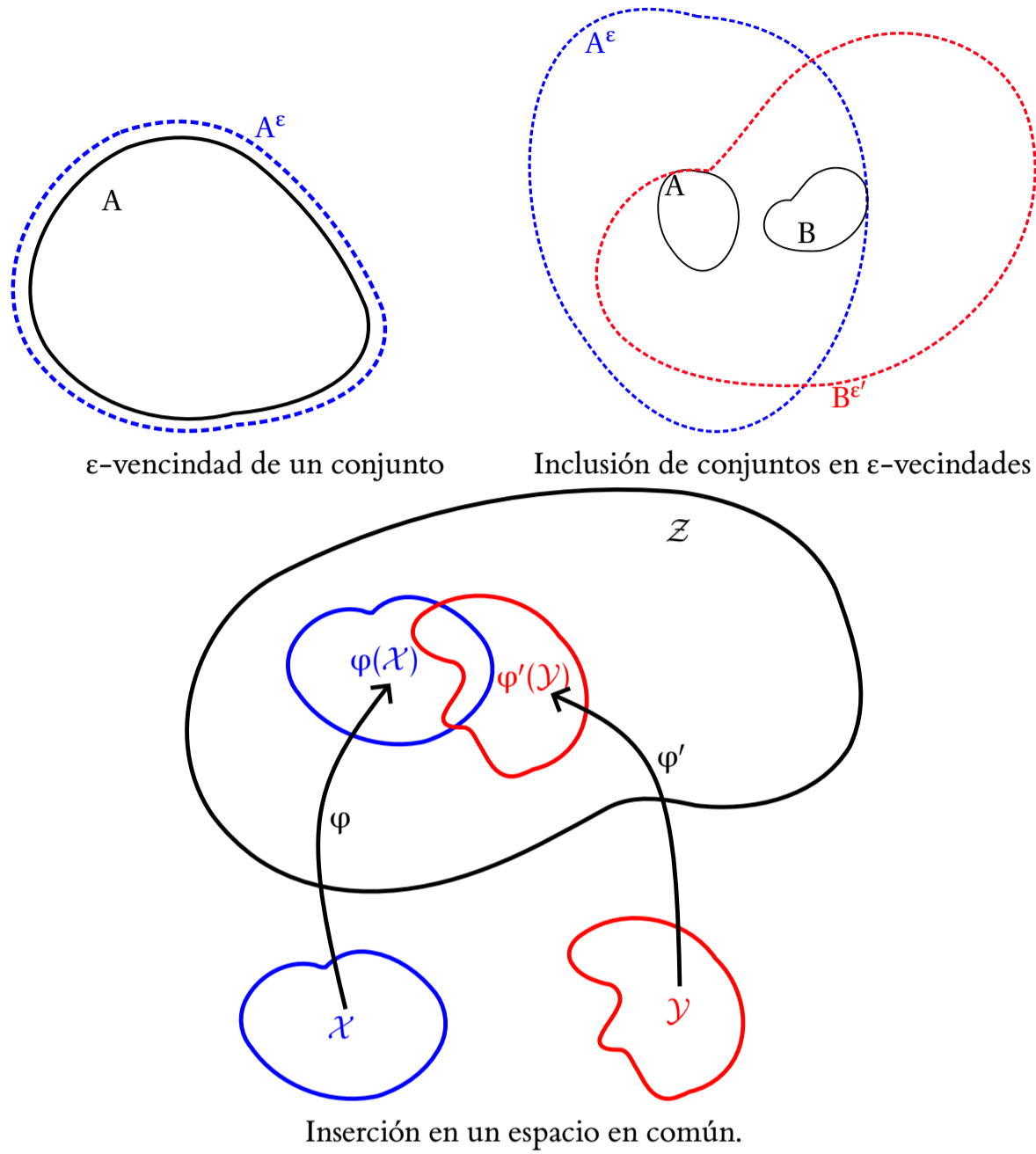


Figura 1: Ingredientes para la distancia de Gromov–Hausdorff.

donde  $\mathcal{C}(\mathcal{X}, \mathcal{Y})$  denota al conjunto de correspondencias entre  $\mathcal{X}$  y  $\mathcal{Y}$ . Esta igualdad se muestra para árboles en el lema 2.3 de Evans et al. (2006); Le Gall afirma que el resultado es también válido para espacios métricos compactos, lo cual se puede verificar en el teorema 7.3.25 de Burago et al. (2001).

Para finalizar la sección, realizaremos la demostración del teorema 2. En la prueba se usa la caracterización de árboles reales a través de una caracterización que dice que un espacio métrico  $(\mathcal{X}, d_{\mathcal{X}})$  es un árbol, si únicamente si, es completo, arco-conexo y satisface la condición de *cuatro puntos*, es decir que para cualesquiera  $x_1, x_2, x_3, x_4 \in \mathcal{X}$ ,

$$d_{\mathcal{X}}(x_1, x_2) + d_{\mathcal{X}}(x_3, x_4) \leq \max\{d_{\mathcal{X}}(x_1, x_3) + d_{\mathcal{X}}(x_2, x_4), d_{\mathcal{X}}(x_1, x_4) + d_{\mathcal{X}}(x_2, x_3)\}.$$

Esta afirmación se puede verificar en el lema II.4.13 del libro de Chiswell (2001). En la demostración usaremos el siguiente lema auxiliar.

### 3. Lema

*El conjunto  $\mathbb{T}$  de árboles reales es un subconjunto cerrado del espacio de espacios métricos*

*compactos equipado con la métrica de Gromov–Hausdorff.*

Demostración: Por el teorema 7.5.1 de Burago et al. (2001) sabemos que el límite de una sucesión en  $\mathbb{T}$  es arco-conexo, mientras que la discusión posterior a la proposición 7.4.12 de Burago et al. (2001) implica que el límite satisface la condición de cuatro puntos. ■

Demostración del teorema 2: Comenzamos con la separabilidad. Dados un árbol  $\mathcal{T} \in \mathbb{T}$  y  $\varepsilon > 0$ , sea  $S_\varepsilon$  una  $\varepsilon$ -red finita en  $\mathcal{T}$ . Denotaremos por  $\mathcal{T}_\varepsilon$  al *subárbol de  $\mathcal{T}$  generado por  $S_\varepsilon$* , es decir

$$\mathcal{T}_\varepsilon := \bigcup_{x,y \in S_\varepsilon} \llbracket x, y \rrbracket \quad \text{con distancia } d_\varepsilon := d|_{\mathcal{T}_\varepsilon}.$$

Por construcción  $\mathcal{T}_\varepsilon$  sigue siendo una  $\varepsilon$ -red para  $\mathcal{T}$ , por lo cual se obtienen las cotas  $d_{\text{GH}}(\mathcal{T}_\varepsilon, \mathcal{T}) \leq d_{\text{H}}(\mathcal{T}_\varepsilon, \mathcal{T}) \leq \varepsilon$ .

De la construcción realizada,  $\mathcal{T}_\varepsilon$  es un “árbol discreto con longitudes para las aristas”, el cual puede ser aproximado bajo  $d_{\text{GH}}$  por árboles con misma topología y longitudes de aristas racionales. El conjunto de árboles discretos con longitudes racionales de aristas es contable, lo cual muestra la separabilidad de  $(\mathbb{T}, d_{\text{GH}})$ .

Resta demostrar la completud del espacio. Gracias al lema 3, bastará probar que cualquier sucesión de Cauchy en  $\mathbb{T}$  converge a un espacio métrico compacto, lo cual es equivalente a que cualquier sucesión de Cauchy en  $\mathbb{T}$  tenga una subsucesión que converja a un espacio métrico compacto.

Sea pues  $\{\mathcal{T}_n : n \in \mathbb{N}\} \subset \mathbb{T}$  una sucesión de Cauchy. En vista de la definición 7.4.13 y el teorema 7.4.15 de Burago et al. (2001), una condición suficiente para que la sucesión sea relativamente compacta y tenga una subsucesión convergente, es que para cada  $\varepsilon > 0$  exista  $N \in \mathbb{N}$  tal que todo elemento  $\mathcal{T}_n$  contenga una  $\varepsilon$ -red de al menos  $N$  puntos.

Fijemos  $\varepsilon > 0$  y  $n_0 \in \mathbb{N}$  tal que  $d_{\text{GH}}(\mathcal{T}_n, \mathcal{T}_m) < \varepsilon/4$  siempre que  $m, n \geq n_0$ . Sea  $S_{n_0}$  una  $(\varepsilon/2)$ -red para  $\mathcal{T}_{n_0}$  de cardinalidad  $N$ . Debido a la caracterización (1) de la distancia de Gromov–Hausdorff, para cada  $n \geq n_0$  existe una correspondencia  $\mathcal{R}_n$  entre  $\mathcal{T}_{n_0}$  y  $\mathcal{T}_n$  que satisface la desigualdad  $\text{dis}(\mathcal{R}_n) < \varepsilon/2$ . Para cada  $x \in \mathcal{T}_{n_0}$  escojamos  $f_n(x) \in \mathcal{T}_n$  de tal forma que  $(x, f_n(x)) \in \mathcal{R}_n$ . Afirmamos que  $S_n = f_n(S_{n_0})$  es una  $\varepsilon$ -red para  $\mathcal{T}_n$  cuando  $n \geq n_0$ . Ciertamente, dado  $y \in \mathcal{T}_n$ , existe  $x \in \mathcal{T}_{n_0}$  tal que  $(x, y) \in \mathcal{R}_n$ . Por definición de  $(\varepsilon/2)$ -red, existe  $x' \in S_{n_0}$  tal que  $d_{\mathcal{T}_{n_0}}(x, x') < \varepsilon/2$ , que junto a la definición de distorsión y que  $(x', f_n(x')) \in \mathcal{R}_n$  implica

$$d_{\mathcal{T}_n}(y, f_n(x')) \leq \text{dis}(\mathcal{R}_n) + d_{\mathcal{T}_{n_0}}(x, x') < \varepsilon.$$

Con ello hemos verificado la completud de  $(\mathbb{T}, d_{\text{GH}})$ . ■

## 5.2. Codificación de árboles reales

Habiendo dado la definición de árboles que usaremos de ahora en adelante, en lo que sigue nos enfocaremos en describir un método para construir árboles reales. Este se adaptará bien a nuestras aplicaciones sobre árboles aleatorios. Sea  $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  una función continua con soporte compacto y  $g(0) = 0$ . Para evitar trivialidades también supondremos que  $g$  no es idénticamente cero. Dados  $s, t \geq 0$  definimos

$$m_g(s, t) = \inf_{r \in [s \wedge t, s \vee t]} g(r)$$

y

$$d_g(s, t) = g(s) + g(t) - 2m_g(s, t).$$

Es claro que  $d_g$  es una función simétrica, es decir  $d_g(s, t) = d_g(t, s)$ , y además es sencillo verificar que se satisface la desigualdad del triángulo, es decir

$$d_g(s, u) \leq d_g(s, t) + d_g(t, u)$$

para cualesquiera  $s, t, u \geq 0$ . Introducimos la relación de equivalencia  $s \sim t$  si, solamente si,  $d_g(s, t) = 0$ ; esto es lo mismo que  $g(s) = g(t) = m_g(s, t)$ . Sea  $\mathcal{T}_g$  es espacio cociente inducido por  $\sim_g$ ,

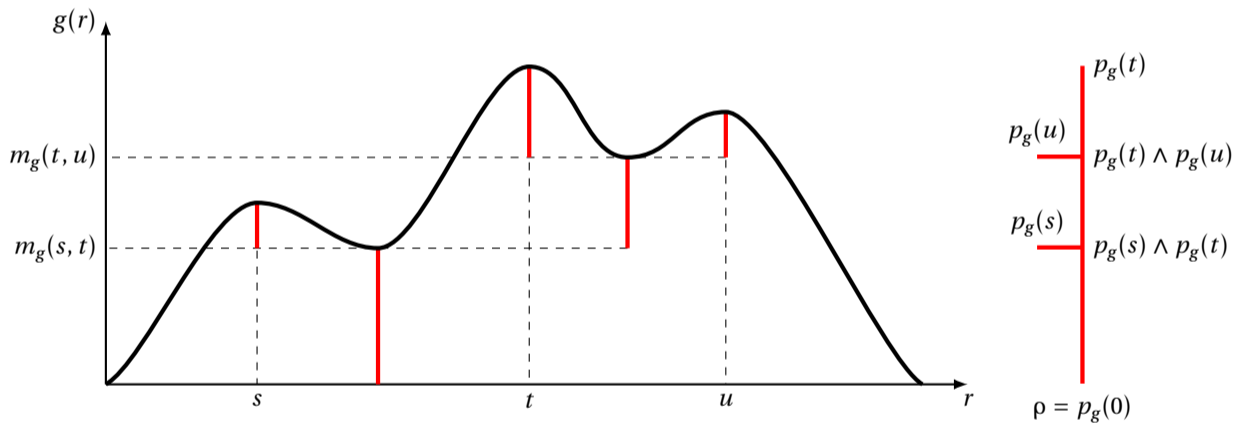
$$\mathcal{T}_g = \mathbb{R}_+ / \sim_g.$$

Por construcción,  $d_g$  induce una métrica sobre  $\mathcal{T}_g$  a la cual también denotaremos por  $d_g$ . Denotamos por  $p_g : \mathbb{R}_+ \rightarrow \mathcal{T}_g$  a la proyección canónica. Si equipamos a  $\mathbb{R}_+$  con la métrica euclidiana y a  $\mathcal{T}_g$  con  $d_g$ , es claro que  $p_g$  es una función continua.

En lo que sigue  $\rho = p_g(0)$ . Además, ponemos  $p_g(t) = \rho$  para cada  $t \geq \zeta$ , donde  $\zeta := \sup(\text{supp } g)$ . Así,  $\mathcal{T}_g = p_g([0, \zeta])$  es compacto.

### 1. Teorema

*El espacio métrico  $(\mathcal{T}_g, d_g)$  es un árbol real, al cual veremos como un árbol enraizado con raíz  $\rho = p_g(0)$ .*



Función  $g$  y rastreo de ancestros comunes más recientes.

Árbol ancestral.

Figura 2: Ancestría de los vértices  $p_g(s)$ ,  $p_g(t)$  y  $p_g(u)$ .

En la figura 2 se muestra como construir un subárbol de  $\mathcal{T}_g$ , el árbol reducido que consiste en la unión de las líneas ancestrales en  $\mathcal{T}_g$  de los vértices  $p_g(s)$ ,  $p_g(t)$ ,  $p_g(u)$  que corresponden a los tiempos  $s, t, u, \in [0, \zeta]$ . El árbol reducido es la unión de las líneas rojas que se construyen de la gráfica de  $g$  como se muestra en la parte izquierda de la figura. La longitud de las líneas discontinuas es despreciable en la construcción y el árbol reducido debe entenderse como en la parte derecha de la figura 2. La línea ancestral de  $p_g(s)$  es un segmento de línea de longitud  $g(u)$ . Asimismo, las líneas ancestrales de  $p_g(s)$  y  $p_g(t)$  tienen una parte en común, la cual tienen longitud  $m_g(s, t)$ .

Más adelante presentamos una demostración elemental del teorema 1 que únicamente usa la definición de árbol real y que ayuda en el entendimiento de línea ancestral y el ancestro más común en  $\mathcal{T}_g$ . Un argumento alternativo, basado en el

teorema 1.2 se da al final de la sección. También es posible dar una prueba basada en la condición de *cuatro puntos*. En efecto, tal prueba puede consultarse en la sección 3 del artículo de Evans y Winter (2006).

Previo a presentar la prueba del teorema 1, enunciamos y demostramos un lema de cambio de raíz.

## 2. Lema

Sea  $s_0 \in [0, \zeta)$ . Para cualquier real  $r \in \mathbb{R}_+$  denotamos por  $\bar{r}$  al único elemento de  $[0, \zeta)$  tal que  $r - \bar{r}$  es un múltiplo entero de  $\zeta$ . Definamos  $\tilde{g} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  mediante

$$\tilde{g}(s) = \begin{cases} g(s_0) + g(\overline{s_0 + s}) - 2m_g(s_0, \overline{s_0 + s}) & \text{si } s \in [0, \zeta], \\ 0 & \text{si } s > \zeta. \end{cases}$$

Entonces  $\tilde{g}$  es continua, tiene soporte compacto y satisface  $\tilde{g}(0) = 0$ , de modo que podemos definir  $\mathcal{T}_{\tilde{g}}$ . En adición, para  $s, t \in [0, \zeta]$  se tiene

$$(1) \quad d_{\tilde{g}}(s, t) = d_g(\overline{s_0 + s}, \overline{s_0 + t}).$$

Finalmente, existe una única isometría  $R$  de  $\mathcal{T}_{\tilde{g}}$  sobre  $\mathcal{T}_g$  tal que, para cualquier  $s \in [0, \zeta]$ ,

$$(2) \quad R(p_{\tilde{g}}(s)) = p_g(\overline{s_0 + s}).$$

Suponiendo que el teorema 1 se ha demostrado, entonces vemos que  $\mathcal{T}_{\tilde{g}}$  coincide con el árbol real  $\mathcal{T}_g$  re-enraizado en  $p_g(s_0)$ . Por ello el lema 2 nos dice qué función codifica el árbol  $\mathcal{T}_g$  re-enraizado en cualquier vértice.

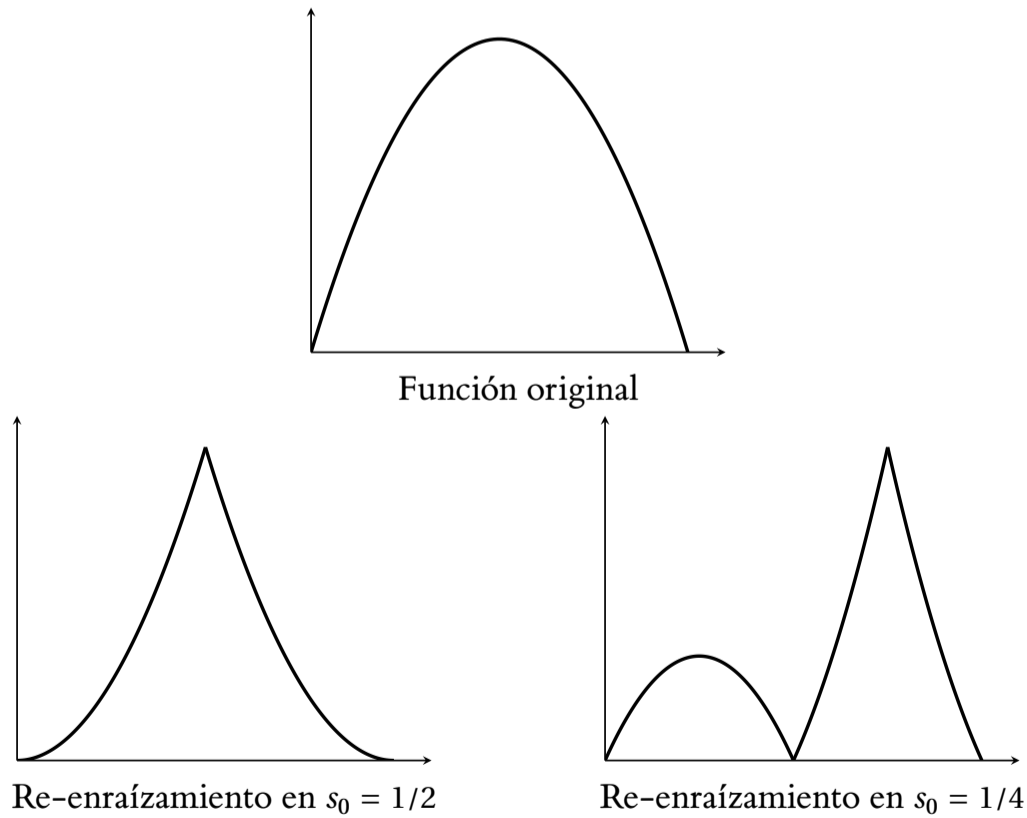


Figura 3: Ejemplo de la función  $\tilde{g}$  con  $g(x) = x(1 - x)$ .

Demostración: Es claro que la función  $\tilde{g}$  satisface los mismos supuestos que  $g$ , por lo cual tiene sentido  $\mathcal{T}_{\tilde{g}}$ . Luego, el paso clave será verificar que se cumple (1). Primero

supondremos que  $s, t \in [0, \zeta - s_0)$ , con  $s \leq t$ , de tal forma que  $\overline{s_0 + s} = s_0 + s$  y  $\overline{s_0 + t} = s_0 + t$ . Veremos que pueden ocurrir dos posibilidades.

Si  $m_g(s_0 + s, s_0 + t) \geq m_g(s_0, s_0 + s)$ , entonces

$$m_g(s_0, s_0 + r) = m_g(s_0, s_0 + s) = m_g(s_0, s_0 + t)$$

para toda  $r \in [s, t]$  y en consecuencia

$$\begin{aligned} m_{\tilde{g}}(s, t) &= g(s_0) - 2m_g(s_0, s_0 + s) + \inf_{r \in [s, t]} g(s_0 + r) \\ &= g(s_0) + m_g(s_0 + s, s_0 + t) - 2m_g(s_0, s_0 + s). \end{aligned}$$

De lo anterior obtenemos

$$\begin{aligned} d_{\tilde{g}}(s, t) &= \tilde{g}(s) + \tilde{g}(t) - 2m_{\tilde{g}}(s, t) \\ &= g(s_0 + s) - 2m_g(s_0, s_0 + s) + g(s_0 + t) - 2m_g(s_0, s_0 + t) \\ &\quad - 2m_g(s_0 + s, s_0 + t) + 4m_g(s_0, s_0 + s) \\ &= g(s_0 + s) + g(s_0 + t) - 2m_g(s_0 + s, s_0 + t) \\ &= d_g(s_0 + s, s_0 + t). \end{aligned}$$

Si en cambio  $m_g(s_0 + s, s_0 + t) < m_g(s_0, s_0 + s)$ , entonces afirmamos que la función  $\tilde{g}$  alcanza su mínimo dentro del intervalo  $[s, t]$  en

$$r_1 = \inf\{r \in [s, t] : g(s_0 + r)\} = m_g(s_0, s_0 + s).$$

Observemos que para  $r_1$  se cumplen tanto  $m_g(s_0, s_0 + s) = m_g(s_0, s_0 + r_1)$  como  $\tilde{g}(r_1) = g(s_0) - m_g(s_0, s_0 + s)$ . Con ello verificaremos que en verdad se alcanza el mínimo en tal punto. En efecto, para  $r \in [s, r_1]$  tendremos que  $m_g(s_0, s_0 + r) = m_g(s_0, s_0 + s)$  y  $g(s_0 + r) \geq g(s_0 + r_1)$ , de modo que

$$\begin{aligned} \tilde{g}(r) &= g(s_0) + g(s_0 + r) - 2m_g(s_0, s_0 + r) \\ &\geq g(s_0) + g(s_0 + r_1) - 2m_g(s_0, s_0 + r_1) = \tilde{g}(r_1). \end{aligned}$$

Mientras tanto, para  $r \in [r_1, t]$  observamos que  $m_g(s_0, s_0 + r) \leq m_g(s_0, s_0 + s)$ , y por consiguiente

$$\begin{aligned} \tilde{g}(r_1) &= g(s_0) - m_g(s_0, s_0 + s) \\ &\leq g(s_0) - m_g(s_0, s_0 + r) \\ &\leq g(s_0) + g(s_0 + r) - 2m_g(s_0, s_0 + r) = \tilde{g}(r). \end{aligned}$$

Así, hemos verificado que se alcanza el mínimo en  $r_1$ ; por ende,

$$m_{\tilde{g}}(s, t) = g(s_0) - m_g(s_0, s_0 + s)$$

y

$$\begin{aligned} d_{\tilde{g}}(s, t) &= g(s_0 + s) - 2m_g(s_0, s_0 + s) + g(s_0 + t) - 2m_g(s_0, s_0 + t) \\ &\quad + 2m_g(s_0, s_0 + s) \\ &= g(s_0 + s) + g(s_0 + t) - 2m_g(s_0 + s, s_0 + t) \\ &= d_g(s_0 + s, s_0 + t). \end{aligned}$$

Los demás casos son similares.

Usando (1), si  $s, t \in [0, \zeta]$  son tales que  $d_{\tilde{g}}(s, t) = 0$ , entonces  $d_g(\overline{s_0 + s}, \overline{s_0 + t}) = 0$  y por lo tanto  $p_g(\overline{s_0 + s}) = p_g(\overline{s_0 + t})$ . Puesto que  $\mathcal{T}_{\tilde{g}} = p_{\tilde{g}}([0, \zeta])$ , es posible definir  $R$  de única manera mediante (2). Para concluir resta notar que (1) implica que  $R$  es una isometría, lo cual implica que  $R$  es sobreyectiva. ■

Demostración del teorema 1: Antes de comenzar con la demostración en sí, establezcamos notación que usaremos y que coincidirá con la notación dada previamente para los árboles. Primero pondremos  $\sigma \leq \sigma'$  si y sólo si  $d_g(\sigma, \sigma') = d_g(\rho, \sigma') - d_g(\rho, \sigma)$ . Así, si  $\sigma = p_g(s)$  y  $\sigma' = p_g(t)$ , notamos que  $\sigma \leq \sigma'$  precisamente cuando  $m_g(s, t) = g(s)$ . Es inmediato que  $\leq$  define un orden parcial sobre  $\mathcal{T}_g$ .

Dados  $\sigma_0, \sigma \in \mathcal{T}_g$ , definimos

$$[\sigma_0, \sigma] := \{\sigma' \in \mathcal{T}_g : d_g(\sigma_0, \sigma) = d_g(\sigma_0, \sigma') + d_g(\sigma', \sigma)\}.$$

De esta definición es claro que si  $\sigma = p_g(s)$  y  $\sigma' = p_g(t)$  entonces  $[\rho, \sigma] \cap [\rho, \sigma'] = [\rho, \gamma]$ , donde  $\gamma = p_g(r)$ , con  $r$  siendo cualquier tiempo en el cual  $g$  alcance su mínimo entre  $s$  y  $t$ . Entonces pondremos  $\gamma := \sigma \wedge \sigma'$ , lo cual representará al ancestro en común más reciente de  $\sigma$  y  $\sigma'$  en cuanto se pruebe el teorema.

Adicionalmente consideramos  $\mathcal{T}_g[\sigma] := \{\sigma' \in \mathcal{T}_g : \sigma \leq \sigma'\}$ , que coincidirá con el subárbol de  $\mathcal{T}_g$  cuya raíz es  $\sigma$ . En caso de tener  $\mathcal{T}_g[\sigma] \neq \{\sigma\}$  y  $\sigma \neq \rho$ , entonces  $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$  y  $\mathcal{T}_g[\sigma] \setminus \{\sigma\}$  resultan ser dos conjuntos abiertos disjuntos. Para ver que  $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$  es, en efecto, abierto, tomemos  $s$  de tal forma que  $p_g(s) = \sigma$  y notemos que  $\mathcal{T}_g[\sigma]$  es la imagen del compacto  $\{u \in [0, \zeta] : m_g(s, u) = g(s)\}$  bajo el mapeo continuo  $p_g$ . Por su parte,  $\mathcal{T}_g[\sigma] \setminus \{\sigma\}$  resulta ser abierto pues si  $\sigma' \in \mathcal{T}_g[\sigma] \setminus \{\sigma\}$ , entonces la bola abierta centrada en  $\sigma'$  y con radio  $d_g(\sigma, \sigma')$  se encuentra contenida en  $\mathcal{T}_g[\sigma] \setminus \{\sigma\}$ .

Dicho lo anterior, probaremos que  $(\mathcal{T}_g, d_g)$  satisface la propiedad (i) de un árbol real. Para ello demostraremos que para  $\sigma_1, \sigma_2 \in \mathcal{T}_g$  dados existe un único mapeo  $f_{\sigma_1, \sigma_2}$  con la propiedad deseada. En vista del lema 2 basta considerar  $\sigma_1 = \rho$ . Así, si  $\sigma \in \mathcal{T}_g$  es dado, probaremos que existe una única isometría  $f = f_{\rho, \sigma}$  de  $[0, d_g(\rho, \sigma)]$  en  $\mathcal{T}_g$  de tal forma que  $f(0) = \rho$  y  $f(d_g(\rho, \sigma)) = \sigma$ . Tomemos  $s \in p_g^{-1}(\{\sigma\})$  de modo que  $g(s) = d_g(\rho, \sigma)$  por definición. Luego, para cada  $a \in [0, d_g(\rho, \sigma)]$  definamos

$$v(a) = \inf\{r \in [0, s] : m_g(r, s) = a\}.$$

Por construcción  $g(v(a)) = a$ , de tal forma que pondremos  $f(a) = p_g(v(a))$ . De esta manera obtendremos directamente  $f(0) = \rho$ , mientras que  $f(d_g(\rho, \sigma)) = \sigma$  se deducirá de que  $m_g(v(g(s)), s) = g(s)$  implica  $p_g(v(g(s))) = p_g(s) = \sigma$ . En adición, dados  $a, b \in [0, d_g(\rho, \sigma)]$  con  $a \leq b$ , es claro que  $m_g(v(a), v(b)) = a$  y por ende

$$d_g(f(a), f(b)) = g(v(a)) + g(v(b)) - 2a = b - a;$$

es decir que  $f$  define una isometría.

Ahora verificaremos la unicidad de  $f$ . Supongamos pues que  $\tilde{f}$  es una isometría que satisface las mismas condiciones que  $f$ . Entonces, para  $a \in [0, d_g(\rho, \sigma)]$ ,

$$d_g(\tilde{f}(a), \sigma) = d_g(\rho, \sigma) - a = d_g(\rho, \sigma) - d_g(\rho, \tilde{f}(a)),$$

es decir que  $\tilde{f}(a) \leq \sigma$ . Recordando que  $\sigma = p_g(s)$  y escogiendo  $t$  de tal manera que  $p_g(t) = \tilde{f}(a)$  observamos que  $g(t) = d_g(\rho, p_g(t)) = a$ . Así, como  $\tilde{f}(a) \leq \sigma$  entonces  $g(t) = m_g(t, s)$ . Por otra parte sabemos que  $a = g(v(a)) = m_g(v(a), s)$ . De ambas afirmaciones se sigue que  $a = g(t) = g(v(a)) = m_g(v(a), t)$  y en consecuencia  $d_g(t, v(a)) = 0$ . Por lo tanto  $\tilde{f}(a) = p_g(t) = p_g(v(a)) = f(a)$ , lo cual muestra la unicidad de  $f$ .

Antes de demostrar que se tiene la segunda propiedad de árbol real, notemos que el argumento previamente expuesto implica  $f([0, d_g(\rho, \sigma)]) = [\rho, \sigma]$ . Ciertamente,

hemos visto que para cada  $a \in [0, d_g(\rho, \sigma)]$ ,  $f(a) \leq \sigma$  y, por otro lado, si  $\eta \leq \sigma$  entonces  $\eta = f(d_g(\rho, \eta))$ .

Dicho lo anterior pasemos a verificar la propiedad (ii) de un árbol real. Sea entonces  $q$  una función continua e inyectiva de  $[0, 1]$  a  $\mathcal{T}_g$ . Demostraremos que  $q([0, 1]) = f_{q(0), q(1)}([0, d_g(q(0), q(1))])$ . Nuevamente podemos suponer  $q(0) = \rho$  gracias al lema 2. Poniendo  $\sigma = q(1)$ , lo expuesto con anterioridad nos da la igualdad  $f_{\rho, \sigma}([0, d_g(\rho, \sigma)]) = [\rho, \sigma]$ . Que  $[\rho, \sigma] = q([0, 1])$  se cumple se demostrará por doble contención, usando argumentos por contradicción en ambos sentidos.

Comenzamos con  $[\rho, \sigma] \subset q([0, 1])$ . Supongamos que existe  $\eta \in [\rho, \sigma] \setminus q([0, 1])$  y en particular  $\eta \notin \{\rho, \sigma\}$ . Bajo este supuesto  $q([0, 1]) \subset (\mathcal{T}_g \setminus \mathcal{T}_g[\eta]) \cup (\mathcal{T}_g[\eta] \setminus \{\eta\})$  con  $q(0) = \rho \in \mathcal{T}_g \setminus \mathcal{T}_g[\eta]$  y  $q(1) = \sigma \in \mathcal{T}_g[\eta] \setminus \{\eta\}$ . Puesto que la unión es de abiertos mutuamente ajenos, esto contradice el hecho de que  $q([0, 1])$  es conexo por ser la imagen continua de un conjunto conexo.

Ahora probaremos que  $q([0, 1]) \subset [\rho, \sigma]$ , nuevamente por contradicción. Así, supongamos que existe  $a \in (0, 1)$  tal que  $q(a) \notin [\rho, \sigma]$ . Definamos  $\eta = q(a)$  y pongamos  $\gamma = \sigma \wedge \eta$ . Entonces  $\gamma \in [\eta, \sigma]$  puesto que  $d_g(\eta, \sigma) = d_g(\eta, \gamma) + d_g(\gamma, \sigma)$  y por ende  $\gamma \in [\rho, \eta] \cap [\eta, \sigma]$ . Esto en particular implica  $\gamma \in q([0, a])$  y  $\gamma \in q([a, 1])$ , lo cual es únicamente posible si  $\gamma = q(a) = \eta$  debido a la inyectividad de  $q$ . No obstante, esto contradice que  $\eta \notin [\rho, \sigma]$ . ■

Ahora que sabemos que, en efecto,  $(\mathcal{T}_g, d_g)$  es un árbol real, la notación dada en la demostración previa,  $\sigma \leq \sigma'$ ,  $[\sigma, \sigma']$  y  $\sigma \wedge \sigma'$ , es consistente con aquella dada para árboles reales en general.

Antes de continuar cabe observar que para un árbol real enraizado  $(\mathcal{T}, d)$  es posible encontrar una función continua  $h \in C_c(\mathbb{R}_+, \mathcal{T})$  de tal forma que recorra el contorno de  $\mathcal{T}$ . Entonces al definir  $g(t) = d(\rho, h(t))$  obtenemos una función  $g$  para la cual  $(\mathcal{T}, d) = (\mathcal{T}_g, d_g)$ ; es decir que todo árbol real enraizado se puede describir de esta manera.

La construcción realizada nos permite discutir sobre las multiplicidades de los vértices del árbol  $\mathcal{T}_g$ . En efecto, si  $\sigma \in \mathcal{T}_g$  no es una hoja, entonces tendremos que  $l(\sigma) < r(\sigma)$ , donde

$$l(\sigma) := \inf p_g^{-1}(\{\sigma\}) \quad \text{y} \quad r(\sigma) := \sup p_g^{-1}(\{\sigma\})$$

son, respectivamente, el menor y mayor elemento de la clase de equivalencia de  $\sigma$  en  $[0, \zeta]$ . Dado lo anterior, tenemos  $m_g(l(\sigma), r(\sigma)) = g(l(\sigma)) = g(r(\sigma)) = d_g(\rho, \sigma)$ . Denotemos por  $(a_i, b_i)$  a las componentes conexas del conjunto abierto

$$(l(\sigma), r(\sigma)) \cap \{t \in \mathbb{R}_+ : g(t) > d_g(\rho, \sigma)\}$$

con  $i \in \mathcal{I}$ , donde  $\mathcal{I}$  es vacío si  $\sigma$  es una hoja. Afirmamos que las componentes conexas del conjunto abierto  $\mathcal{T}_g \setminus \{\sigma\}$  son precisamente  $p_g((a_i, b_i))$ ,  $i \in \mathcal{I}$  y, cuando  $\sigma \neq \rho$ ,  $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$ , de modo que la multiplicidad de  $\rho$  queda dada por  $1 + \#\mathcal{I}$ . Ya se estableció que  $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$  es abierto, mientras que el argumento dado para  $\mathcal{T}_g[\sigma] \setminus \{\sigma\}$  también nos dice que  $p_g((a_i, b_i))$  es abierto. Ahora, por tratarse de imágenes continuas de intervalos, los conjuntos  $p_g((a_i, b_i))$  son conexos, mientras que  $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$  es conexo pues para  $\sigma', \sigma'' \in \mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$  se tiene que  $[\rho, \sigma] \cup [\rho, \sigma'']$  es un conjunto cerrado y conexo contenido en  $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$ .

En el siguiente resultado se comparan los árboles codificados mediante dos funciones  $g$  y  $\tilde{g}$ .

**3. Lema**

Sean  $g, \tilde{g} \in C_c(\mathbb{R}_+, \mathbb{R}_+)$  con  $g(0) = \tilde{g}(0) = 0$ . Entonces

$$d_{\text{GH}}(\mathcal{T}_g, \mathcal{T}_{\tilde{g}}) \leq 2\|g - \tilde{g}\|_\infty.$$

Demostración: En la prueba usaremos la caracterización (1.1) de la distancia de Gromov–Hausdorff. Para ello definiremos una correspondencia  $\mathcal{R}$  entre  $\mathcal{T}_g$  y  $\mathcal{T}_{\tilde{g}}$  mediante

$$\mathcal{R} := \{(\sigma, \sigma') : \sigma = p_g(t) \text{ y } \sigma' = p_{\tilde{g}}(t) \text{ para algún } t \geq 0\}.$$

Con ella, tomemos  $(\sigma, \sigma') \in \mathcal{R}$  y  $(\eta, \eta') \in \mathcal{R}$ , de modo que podemos encontrar  $s, t \geq 0$  tales que  $p_g(s) = \sigma$ ,  $p_{\tilde{g}}(s) = \sigma'$ ,  $p_g(t) = \eta$  y  $p_{\tilde{g}}(t) = \eta'$ . Recordando que

$$d_g(\sigma, \eta) = g(s) + g(t) - 2m_g(s, t)$$

y

$$d_{\tilde{g}}(\sigma', \eta') = \tilde{g}(s) + \tilde{g}(t) - 2m_{\tilde{g}}(s, t),$$

deducimos

$$|d_g(\sigma, \eta) - d_{\tilde{g}}(\sigma', \eta')| \leq 4\|g - \tilde{g}\|_\infty.$$

Lo anterior nos da  $\text{dis}(\mathcal{R}) \leq 4\|g - \tilde{g}\|_\infty$ , desigualdad que nos permite concluir en virtud de (1.1). ■

Usando el lema 3 podemos dar una demostración alternativa del teorema 1. Sea  $C_{c,0}$  el conjunto de funciones  $g \in C_c(\mathbb{R}_+, \mathbb{R}_+)$  con  $g(0) = 0$  y para las cuales se cumple que existen  $\varepsilon > 0$  y  $\rho > 0$  tales que, para cada  $i \in \mathbb{N}$ ,  $g$  es lineal con pendiente  $\rho$ ,  $-\rho$  o 0 en el intervalo  $[(i-1)\varepsilon, i\varepsilon]$ . Entonces  $g$  coincide con la función de contorno de un árbol discreto  $\mathbf{t} \in \mathbf{T}$  debidamente reescalado, dándonos un árbol  $\mathcal{T}_g$ . Luego, como una función  $g \in C_c(\mathbb{R}_+, \mathbb{R}_+)$  con  $g(0) = 0$  puede ser escrita como el límite uniforme de una sucesión  $\{g_n\} \subset C_{c,0}$ , el lema 3 implica que  $\mathcal{T}_g$  es el límite de  $\{\mathcal{T}_{g_n}\}$  bajo la métrica de Gromov–Hausdorff. Así obtenemos que  $\mathcal{T}_g$  es un árbol real.

### 5.3. El árbol aleatorio continuo (CRT)

Recordemos que  $e = \{e_t : t \in [0, 1]\}$  es la excursión browniana normalizada, que extenderemos a  $\mathbb{R}_+$  poniendo  $e_t = 0$  si  $t > 1$ . Comenzamos definiendo el objeto conocido como árbol aleatorio continuo, o CRT por sus siglas en inglés: *Continuum Random Tree*.

**1. Definición**

El árbol aleatorio continuo es el árbol real  $\mathcal{T}_e$  codificado por la excursión browniana normalizada.

De la definición vemos que  $\mathcal{T}_e$  es un elemento aleatorio que toma valores en el espacio  $\mathbb{T}$ , el cual es medible gracias al lema 2.3. De hecho, el mismo lema nos permitirá traducir los resultados de convergencia del capítulo 4 en términos de convergencia débil en el espacio  $\mathbb{T}$ . En este apartado veremos un ejemplo típico de

cómo es que el CRT es el límite de árboles “combinatorios” en lugar de traducir todos los resultados previamente obtenidos.

En el capítulo 1 denotamos por  $\mathbf{T}$  al conjunto de todos los árboles finitos enraizados mientras que en el capítulo 4 denotamos por  $\mathbf{T}_p$  al subconjunto de  $\mathbf{T}$  de árboles con  $p$  vértices. Es posible ver cada elemento  $t \in \mathbf{T}$  como un árbol real enraizado, considerando a  $t$  como la unión de segmentos de longitud 1, como se representa en la figura 1.1.1, considerando la distancia entre  $\sigma$  y  $\sigma'$  como la longitud del camino más corto entre ambos vértices. De manera alternativa podemos considerar el árbol real  $\mathcal{T}_C$  si  $C$  es la función de contorno de  $t$ .

Para cualquier elección de  $\lambda > 0$  y  $\mathcal{T} \in \mathbb{T}$ , denotaremos por  $\lambda\mathcal{T}$  al árbol  $\mathcal{T}$  equipado con la métrica  $\lambda d_{\mathcal{T}}$ , es decir que es el “mismo” árbol pero con las distancias multiplicadas por un factor  $\lambda$ .

## 2. Teorema

Para cada  $n \geq 1$  sea  $\mathcal{T}_{(n)}$  un árbol aleatorio con distribución uniforme sobre  $\mathbf{T}_n$ . Entonces  $(2n)^{-1/2}\mathcal{T}_{(n)}$  converge en distribución al CRT  $\mathcal{T}_e$  en el espacio  $(\mathcal{T}, d_{\text{GH}})$ .

Demostración: Sea  $\theta$  un árbol de Galton–Watson con distribución de reproducción geométrica, es decir  $\mu(k) = 2^{-k-1}$  y para cada  $n$  sea  $\theta^{(n)}$  el árbol condicionado a tener  $n$  vértices. Por las observaciones realizadas en la sección 4.2,  $\theta^{(n)}$  tiene la misma distribución que  $\mathcal{T}_{(n)}$ . Dada la función de contorno  $C^{(n)}$  de  $\theta^{(n)}$ , para cada  $t \geq 0$  definimos

$$\tilde{C}_t^n = \frac{1}{\sqrt{2n}} C_{2nt}^{(n)}.$$

Entonces el teorema IV.2.1 reescrito en términos de la función de contorno implica la convergencia  $\tilde{C}^{(n)} \Rightarrow e$  conforme  $n \rightarrow \infty$ . Resta notar que  $\mathcal{T}_{\tilde{C}^{(n)}}$  tiene la misma distribución que  $(2n)^{-1/2}\mathcal{T}_{(n)}$ , lo cual nos permite concluir gracias al lema 2.3. ■

Es posible considerar resultados análogos al del teorema 2 para los árboles combinatorios considerados al final de la sección 4.2, donde vemos que si  $\tau_n$  tiene distribución uniforme sobre todos los árboles de Cayley con  $n$  vértices, o todos los árboles binarios con  $n$  vértices, entonces  $(4n)^{-1/2}\tau_n$  converge en distribución a  $\mathcal{T}_e$ .

## 5.4. La medida de excursiones de Itô

Es de nuestro interés el derivar explícitamente ciertas distribuciones asociadas al CRT. De manera más específica, nos interesan las distribuciones finito-dimensionales. Para realizar los cálculos usaremos algunas propiedades de las excursiones brownianas, que pueden verse como un caso particular de las excursiones estudiadas en la sección 2.4. Antes de trabajar con la excursión browniana normalizadas, veremos algunos resultados particulares de la medida de excursión de Itô para el movimiento browniano. Cabe destacar que lo que se verá no se sigue directamente de los resultados en la sección 2.4 debido a que aquí nos centraremos en excursiones positivas.

Denotaremos por  $B = \{B_t : t \geq 0\}$  a un movimiento browniano lineal que comienza en  $x$  bajo la medida de probabilidad  $\mathbb{P}_x$ , entendiendo  $\mathbb{P}$  por  $\mathbb{P}_0$ . Definimos

los procesos  $S = \{S_t : t \geq 0\}$  e  $I_t = \{I_t : t \geq 0\}$  mediante

$$S_t = \sup_{s \leq t} B_s \quad \text{y} \quad I_t = \inf_{s \leq t} B_s.$$

En adición, para cada  $a \in \mathbb{R}$  ponemos  $T_a = \inf\{t \geq 0 : B_t = a\}$ . Recordamos que el principio de reflexión nos dice que para  $a \geq 0$  y  $b \leq a$ ,

$$\mathbb{P}(S_t \geq a, B_t \leq b) = \mathbb{P}(B_t \geq 2a - b),$$

es decir que nos da la ley del par  $(S_t, B_t)$ . Al derivar con respecto a  $a$  y  $b$  se deduce que, bajo  $\mathbb{P}$ , la ley del par  $(S_t, B_t)$  tiene densidad dada por

$$\gamma_t(a, b) = \frac{2(2a - b)}{\sqrt{2\pi t^3}} \exp\left\{-\frac{(2a - b)^2}{2t}\right\} \mathbb{1}_{\{a \geq 0, b \leq a\}}.$$

No solo eso, el principio de reflexión también implica que  $S_t$  y  $|B_t|$  tienen la misma distribución.

Ahora, dado  $a > 0$ , es claro que bajo  $\mathbb{P}$ ,  $\{T_a \leq t\} = \{S_t \geq a\}$  casi seguramente. En consecuencia, la distribución de  $T_a$ , bajo  $\mathbb{P}$ , tiene densidad

$$q_a(t) = \frac{a}{\sqrt{2\pi t^3}} \exp\left\{-\frac{a^2}{2t}\right\}.$$

La relación  $\gamma_t(a, b) = 2q_{2a-b}(t)$  es inmediata siempre que  $a \geq 0$  y  $b < a$ .

Dado  $\varepsilon > 0$  sea  $\nu_\varepsilon$  la ley de la primera excursión de  $B$  fuera de cero que alcance el nivel  $\varepsilon$ . Dicho de manera formal, si  $G_\varepsilon = \sup\{s < T_\varepsilon : B_s = 0\}$  y  $D_\varepsilon = \inf\{t > T_\varepsilon : B_t = 0\}$ ,  $\nu_\varepsilon$  describe la ley de  $\{B_{(G_\varepsilon+t) \wedge D_\varepsilon} : t \geq 0\}$ . Así, la medida  $\nu_\varepsilon$  resulta ser una medida de probabilidad sobre  $C(\mathbb{R}_+, \mathbb{R}_+)$  y tiene soporte en  $\mathcal{C}_\varepsilon = \{e \in C(\mathbb{R}_+, \mathbb{R}_+) : \sup_s e(s) \geq \varepsilon\}$ . Un resultado elemental nos dice que al considerar  $0 < \varepsilon < \varepsilon'$ , entonces obtenemos la segunda igualdad en

$$\nu_\varepsilon(\mathcal{C}_{\varepsilon'}) = \mathbb{P}_\varepsilon(T_{\varepsilon'} < T_0) = \frac{\varepsilon}{\varepsilon'}.$$

Esta igualdad nos permite deducir

$$\nu_{\varepsilon'} = \nu_\varepsilon(\cdot \mid \mathcal{C}_{\varepsilon'}) = \frac{\varepsilon'}{\varepsilon} \nu_\varepsilon(\cdot \cap \mathcal{C}_{\varepsilon'}).$$

Luego, si para cada  $\varepsilon > 0$  ponemos

$$n_\varepsilon = \frac{1}{2\varepsilon} \nu_\varepsilon,$$

obtenemos  $n_{\varepsilon'} = n_\varepsilon(\cdot \cap \mathcal{C}_{\varepsilon'})$  para cualesquiera  $0 < \varepsilon < \varepsilon'$ . Esto nos da la pauta para la siguiente definición.

### 1. Definición

La medida  $\sigma$ -finita sobre  $C(\mathbb{R}_+, \mathbb{R}_+)$  definida por

$$n_+ := \lim_{\varepsilon \searrow 0} n_\varepsilon$$

se conoce como la *medida de excursiones positivas de Itô del movimiento browniano lineal*.

Nótese que lo único que hemos realizado es definir la medida de excursiones considerando la altura de ellas en lugar de su tiempo de vida o duración, como se hizo en la sección 2.4, y que en lugar de tener soporte en un subconjunto de las funciones càdlàg,  $n_+$  tiene soporte sobre el conjunto  $\mathcal{E}$  de elementos  $e \in C(\mathbb{R}_+, \mathbb{R}_+)$  con la propiedad de que existe  $\zeta = \zeta(e) > 0$  tal que  $e(t) > 0$  si, únicamente si,  $t \in (0, \zeta)$ . El número  $\zeta$  es precisamente la duración o tiempo de vida de la excursión  $e$ . Por construcción  $n_\varepsilon$  es la restricción de  $n_+$  a  $\mathcal{C}_\varepsilon$  y por la misma construcción,  $n_+(\mathcal{C}_\varepsilon) = 1/2\varepsilon$ . Finalmente, si  $T_\varepsilon(e) := \inf\{t \geq 0 : e(t) = \varepsilon\}$ , la distribución de  $\{e(T_\varepsilon(e) + t) : t \geq 0\}$  bajo  $n_+(\cdot | T_\varepsilon < \infty) = \nu_\varepsilon$  es la misma que la ley de  $\{B_{t \wedge T_0} : t \geq 0\}$  bajo  $\mathbb{P}_\varepsilon$ . Esta afirmación es consecuencia directa de la propiedad fuerte de Markov del movimiento browniano y la construcción de  $n_+$ .

## 2. Proposición

(i) Para cada  $t > 0$  y cada función medible  $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  con  $g(0) = 0$ ,

$$\int_{C(\mathbb{R}_+, \mathbb{R}_+)} g(e(t)) \, dn_+(e) = \int_0^\infty q_x(t) g(x) \, dx.$$

En particular  $n_+(\zeta > t) = n_+(e(t) > 0) = (2\pi t)^{-1/2} < \infty$ . Más aún,

$$n_+\left(\int_0^\infty g(e(t)) \, dt\right) = \int_0^\infty g(x) \, dx.$$

(ii) Dadas  $t > 0$ ,  $\Phi : C([0, t], \mathbb{R}_+) \rightarrow \mathbb{R}_+$  y  $\Psi : C(\mathbb{R}_+, \mathbb{R}_+) \rightarrow \mathbb{R}_+$  medibles,

$$\begin{aligned} & \int_{C(\mathbb{R}_+, \mathbb{R}_+)} \Phi(e(r), 0 \leq r \leq t) \Psi(e(t+r), r \geq 0) \, dn_+(e) \\ &= \int_{C(\mathbb{R}_+, \mathbb{R}_+)} \Phi(e(r), 0 \leq r \leq t) \mathbb{E}_{e(t)}[\Psi(B_{r \wedge T_0}, r \geq 0)] \, dn_+(e). \end{aligned}$$

Demostración: En vista del teorema de convergencia monótona y el teorema de clases monótonas, basta demostrar la primera parte de (i) cuando  $g$  es continua, acotada y existe  $\alpha > 0$  tal que  $g(x) = 0$  si  $x \leq \alpha$ . Bajo estas hipótesis para  $g$ , el teorema de convergencia dominada nos da la primera igualdad en

$$\begin{aligned} \int_{C(\mathbb{R}_+, \mathbb{R}_+)} g(e(t)) \, dn_+(e) &= \lim_{\varepsilon \rightarrow 0} \int_{C(\mathbb{R}_+, \mathbb{R}_+)} g(T_\varepsilon(e) + t) \mathbb{1}_{\{T_\varepsilon(e) < \infty\}} \, dn_+(e) \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \mathbb{E}_\varepsilon[g(B_{t \wedge T_0})]. \end{aligned}$$

Ahora notemos que

$$\begin{aligned} \mathbb{E}_\varepsilon[g(B_{t \wedge T_0})] &= \mathbb{E}_\varepsilon[g(B_t) \mathbb{1}_{\{t \leq T_0\}}] = \mathbb{E}[g(B_t + \varepsilon) \mathbb{1}_{\{t \leq T_{-\varepsilon}\}}] \\ &= \mathbb{E}[g(B_t + \varepsilon) \mathbb{1}_{\{S_t \geq -\varepsilon\}}] = \mathbb{E}[g(\varepsilon - B_t) \mathbb{1}_{\{S_t \leq \varepsilon\}}] \\ &= \int_0^\varepsilon \int_{-\infty}^a g(\varepsilon - b) \gamma_t(a, b) \, db \, da, \end{aligned}$$

donde la primera igualdad se da pues  $g(0) = 0$ , la segunda por un argumento de traslación, la cuarta debido a que  $B$  y  $-B$  tienen la misma distribución bajo  $\mathbb{P}$  y la última al usar la densidad de la ley de  $(S_t, B_t)$ . De ambas relaciones deducimos

$$\int_{C(\mathbb{R}_+, \mathbb{R}_+)} g(e(t)) \, dn_+(e) = \frac{1}{2} \int_{-\infty}^0 g(-b) \gamma_t(0, b) \, db = \frac{1}{2} \int_0^\infty g(x) \gamma_t(0, -x) \, dx,$$

con la segunda igualdad debiéndose al cambio de variable  $x = -b$ . Así, basta usar que  $\gamma_t(0, -x) = 2q_x(t)$  para obtener la primera parte de (i). De esto, al considerar  $g(x) = \mathbb{1}_{(0, \infty)}(x)$  obtenemos

$$n_+(\zeta > t) = \int_0^\infty \frac{x}{\sqrt{2\pi t^3}} e^{-x^2/2t} dx = \frac{1}{\sqrt{2\pi t}}.$$

La última afirmación de (i) se sigue de Fubini y de que  $t \mapsto q_x(t)$  es una densidad; efectivamente, tenemos que

$$n_+\left(\int_0^\infty g(e(t)) dt\right) = \int_0^\infty \int_0^\infty q_x(t)g(x) dx dt = \int_0^\infty g(x) dx.$$

Para demostrar (ii) podemos suponer que  $\Phi$  y  $\Psi$  son continuas y acotadas, nuevamente aplicando el teorema de clases monótonas. Más aún, por el soporte de  $n_+$  podemos suponer que existe  $\alpha \in (0, t)$  tal que se tenga que  $\Psi(\omega(r), 0 \leq r \leq t) = 0$  si  $\omega(\alpha) = 0$ . Así, el teorema de convergencia dominada y la propiedad de Markov en el tiempo  $T_0$  nos dan la cadena de igualdades

$$\begin{aligned} & \int_{C(\mathbb{R}_+, \mathbb{R}_+)} \Phi(e(r), 0 \leq r \leq t) \Psi(e(t+r), r \geq 0) dn_+(e) \\ &= \lim_{\varepsilon \rightarrow 0} \int_{C(\mathbb{R}_+, \mathbb{R}_+)} \Phi(e(T_\varepsilon(e) + r), 0 \leq r \leq t) \Psi(e(T_\varepsilon(e) + t + r), r \geq 0) \mathbb{1}_{\{T_\varepsilon(e) < \infty\}} dn_+(e) \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \mathbb{E}_\varepsilon[\Phi(B_{r \wedge T_0}, 0 \leq r \leq t) \Psi(B_{(t+s) \wedge T_0}, s \geq 0)] \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \mathbb{E}_\varepsilon\left[\Phi(B_{r \wedge T_0}, 0 \leq r \leq t) \mathbb{E}_{B_{t \wedge T_0}}[\Psi(B_{s \wedge T_0}, s \geq 0)]\right] \\ &= \lim_{\varepsilon \rightarrow 0} \int_{C(\mathbb{R}_+, \mathbb{R}_+)} \Phi(e(T_\varepsilon(e) + r), 0 \leq r \leq t) \mathbb{E}_{e(T_\varepsilon(e)+t)}[\Psi(B_{s \wedge T_0}, s \geq 0)] \mathbb{1}_{\{T_\varepsilon(e) < \infty\}} dn_+(e) \\ &= \int_{C(\mathbb{R}_+, \mathbb{R}_+)} \Phi(e(r), 0 \leq r \leq t) \mathbb{E}_{e(t)}[\Psi(B_{r \wedge T_0}, r \geq 0)] dn_+(e). \end{aligned}$$

La propiedad de Markov se ha usado en la tercera igualdad, mientras que convergencia dominada se usa en la primera y última igualdades. La aplicación del teorema de convergencia dominada es válido debido a que los supuestos realizados sobre  $\Phi$  y  $\Psi$  implican la existencia de  $K > 0$  tal que

$$\mathbb{1}_\mathcal{E}(e) \Phi(e(T_\varepsilon(e) + r), 0 \leq r \leq t) \Psi(e(T_\varepsilon(e) + t + r), r \geq 0) \mathbb{1}_{\{T_\varepsilon(e) < \infty\}} \leq K \mathbb{1}_{\{\zeta > \alpha\}},$$

y  $n_+(\zeta > \alpha) < \infty$ . ■

## 5.5. Distribuciones finito-dimensionales bajo la medida de Itô

Dado un árbol real  $(\mathcal{T}, d)$  con raíz  $\rho$ , el subárbol generado por  $\{x_1, \dots, x_p\} \subset \mathcal{T}$  se define como el conjunto

$$\mathcal{T}(x_1, \dots, x_p) = \bigcup_{i=1}^p [\rho, x_i].$$

En este apartado estamos interesados en ver la distribución de los subárboles generados dentro de  $\mathcal{T}_e$  bajo la medida de excursiones estudiada en la sección 5.4.

Es claro que  $\mathcal{T}(x_1, \dots, x_p)$  vuelve a ser un árbol real cuando se le equipa con  $d$ ; no obstante, ahora tiene una estructura discreta: es posible representar al árbol mediante un esqueleto discreto, un árbol discreto con  $p$  hojas etiquetadas, y la colección de distancias de las “ramas”, indexadas por los vértices del esqueleto.

En lo que sigue nos concentraremos en los árboles de la forma  $\mathcal{T}_g$ , es decir que se encuentran codificados por una función apropiada  $g$ . Para introducir las ideas que usaremos consideraremos una función general  $g \in C(\mathbb{R}_+, \mathbb{R}_+)$ . Así, dados  $0 \leq t_1 \leq t_2 \leq \dots \leq t_p$  definiremos el “árbol marcado”

$$\theta(g; t_1, \dots, t_p) = \left( \tau(g; t_1, \dots, t_p), \{h_v : v \in \tau(g; t_1, \dots, t_p)\} \right),$$

donde  $\tau(g; t_1, \dots, t_p) \in \mathbf{T}$  y  $h_v \geq 0$  para cada  $v \in \tau(g; t_1, \dots, t_p)$ . Un ejemplo con  $p = 3$  se muestra en la figura 4. En el ejemplo  $\tau(g; t_1, t_2, t_3) = \{\emptyset, 1, 2, 21, 22\}$  y los números  $\{h_v : v \in \tau(g; t_1, t_2, t_3)\}$  se muestran como las longitudes de los segmentos rojos en la figura.

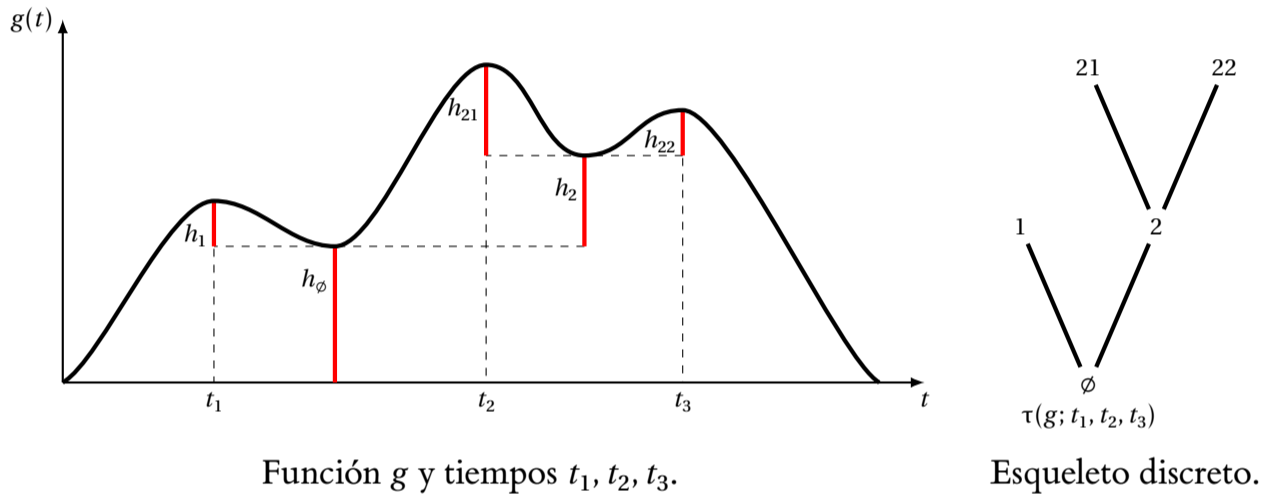


Figura 4: Ejemplo de  $\theta(g; t_1, t_2, t_3)$ .

La definición de  $\theta(g; t_1, \dots, t_p)$  se dará mediante inducción sobre  $p$ :

- Si  $p = 1$  entonces  $\tau(g; t_1) = \{\emptyset\}$  y  $h_{\emptyset}(g; t_1) = g(t_1)$ .
- Para  $p \geq 2$  supondremos que el árbol marcado  $\theta(g; t_1, \dots, t_j)$  se ha construido para  $j \leq p - 1$ . Entonces existen un entero  $k \in \{1, \dots, p - 1\}$  y  $k$  enteros  $1 \leq i_1 < i_2 < \dots < i_k \leq p - 1$  tales que  $m_g(t_i, t_{i+1}) = m_g(t_1, t_p)$  precisamente cuando  $i \in \{i_1, \dots, i_k\}$ . Ponemos  $i_0 = 0$  e  $i_{k+1} = p$  para definir  $g^l \in C(\mathbb{R}_+, \mathbb{R}_+)$ ,  $l \in [k + 1]$ , mediante

$$g^l(t) = g((t \vee t_{i_{l-1}+1}) \wedge t_{i_l}) - m_g(t_1, t_p).$$

Definimos  $\tau(g; t_1, \dots, t_p)$  como la concatenación de los árboles  $\tau(g^l; t_{i_{l-1}+1}, \dots, t_{i_l})$  para  $l \in [k + 1]$ , es decir

$$\tau(g; t_1, \dots, t_p) = \{\emptyset\} \cup \bigcup_{l=1}^{k+1} \{lu : u \in \tau(g^l; t_{i_{l-1}+1}, \dots, t_{i_l})\}.$$

Para construir las marcas, si para  $l \in [k + 1]$ ,

$$\theta(g^l; t_{i_{l-1}+1}, \dots, t_{i_l}) = \left( \tau(g^l; t_{i_{l-1}+1}, \dots, t_{i_l}), \{h_v^l : v \in \tau(g^l; t_{i_{l-1}+1}, \dots, t_{i_l})\} \right),$$

entonces las marcas  $\{h_v : v \in \tau(g; t_1, \dots, t_p)\}$  quedan dadas por  $h_{\emptyset} = m_g(t_1, t_p)$  y

$$h_v = h_u^l \text{ si } v = lu \text{ con } u \in \tau(g^l; t_{i_{l-1}+1}, \dots, t_{i_l}).$$

De esta manera se construye inductivamente el árbol. Puede observarse que  $k + 1$  es el número de hijos de  $\emptyset$  en el árbol  $\theta(g; t_1, \dots, t_p)$ .

Si ahora retornamos al caso que nos concierne, con  $g \in C_c(\mathbb{R}_+, \mathbb{R}_+)$  cumpliendo  $g(0) = 0$ , la construcción de  $\theta(g; t_1, \dots, t_p)$  coincide con el árbol  $\mathcal{T}_g(p_g(t_1), \dots, p_g(t_p))$  generado por los vértices  $p_g(t_1), \dots, p_g(t_p)$  en  $\mathcal{T}_g$ . Ciertamente, si a cada vértice  $v \in \tau(g; t_1, \dots, t_p)$  se le engancha un segmento de longitud  $h_v$  en el plano, de forma que este segmento tenga un “fin” común con los segmentos enganchados a los hijos y que fuera de éste los segmentos no se intersequen, entonces la unión de tales segmentos da un representante de la clase de equivalencia de  $\mathcal{T}_g(p_g(t_1), \dots, p_g(t_p))$ . Un ejemplo de esta construcción para el árbol obtenido en la figura 4 se puede ver en la figura 2.2, identificando  $\rho$  con  $\emptyset$ ,  $p_g(s)$  con 1,  $p_g(t) \wedge p_g(u)$  con 2,  $p_g(t)$  con 21 y  $p_g(u)$  con 22.

Para los resultados que siguen es necesario establecer notación que será de utilidad. Denotaremos por  $\mathbf{T}_{(p)}$  al conjunto de árboles enraizados y ordenados con  $p$  hojas. Por su parte,  $\Theta_{(p)}$  representará al conjunto de árboles marcados con  $p$  hojas, es decir elementos de la forma  $\theta = (\tau, \{h_v : v \in \tau\})$  donde  $\tau \in \mathbf{T}_{(p)}$  y  $h_v \geq 0$  para cada  $v \in \tau$ . Adicionalmente consideraremos los conjuntos de árboles binarios:  $\mathbf{T}_{(p)}^{\text{bin}}$  es el conjunto de árboles binarios enraizados con  $p$  hojas y  $\Theta_{(p)}^{\text{bin}}$  es el conjunto de árboles marcados  $\theta = (\tau, \{h_v : v \in \tau\})$  con esqueleto  $\tau \in \mathbf{T}_{(p)}^{\text{bin}}$ . Es sabido que el número de árboles binarios enraizados con  $p$  hojas es el  $p - 1$  número de Catalan, es decir

$$\#\mathbf{T}_{(p)}^{\text{bin}} = c_{p-1} = \frac{1}{p} \binom{2(p-1)}{p-1} = \frac{(2p-2)!}{(p-1)!p!}.$$

### 1. Teorema

*La ley del árbol  $\theta(e; t_1, \dots, t_p)$  bajo la medida*

$$dn_+(e) \mathbb{1}_{\{0 \leq t_1 \leq \dots \leq t_p \leq \zeta(e)\}} dt_1 \cdots dt_p$$

*es  $2^{p-1} \Lambda_p$ , donde  $\Lambda_p$  es la medida uniforme sobre  $\Theta_{(p)}^{\text{bin}}$  definido mediante*

$$\int_{\Theta_{(p)}^{\text{bin}}} F(\theta) d\Lambda_p(\theta) = \sum_{\tau \in \mathbf{T}_{(p)}^{\text{bin}}} \int_{\mathbb{R}_+^{2p-1}} F(\tau, \{h_v : v \in \tau\}) \prod_{v \in \tau} dh_v.$$

La construcción del árbol marcado  $\theta(e; t_1, \dots, t_p)$  únicamente depende de los valores  $e(t_1), \dots, e(t_p)$  y la colección de mínimos  $m_e(t_1, t_2), \dots, m_e(t_{p-1}, t_p)$ . Por ello el siguiente resultado será esencial en la demostración del teorema 1. Con aras de reducir la notación usaremos  $m$  en lugar de  $m_e$ .

### 2. Proposición

*Sea  $F$  no negativa y medible sobre  $\mathbb{R}_+^{2p-1}$ . Se cumple*

$$\begin{aligned} n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \zeta\}} F(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p)) dt_1 \cdots dt_p \right) \\ = 2^{p-1} \int_{\mathbb{R}_+^{2p-1}} \left( \prod_{i=1}^{p-1} \mathbb{1}_{\{\alpha_i \leq \beta_i \wedge \beta_{i+1}\}} \right) F(\alpha_1, \dots, \alpha_{p-1}, \beta_1, \dots, \beta_p) d\alpha_1 \cdots d\alpha_{p-1} d\beta_1 \cdots d\beta_p. \end{aligned}$$

Antes de probar la proposición 2, enunciamos y demostramos un lema que permitirá simplificar la demostración de la proposición 2.

**3. Lema**

Si  $g : \mathbb{R}^3 \rightarrow \mathbb{R}_+$  es una función medible y  $x \geq 0$ , entonces

$$\mathbb{E}_x \left[ \int_0^{T_0} g(t, I_t, B_t) dt \right] = 2 \int_0^x \int_y^\infty \int_0^\infty q_{x+z-2y}(t) g(t, y, z) dt dz dy.$$

En particular, dada una función  $h : \mathbb{R}^2 \rightarrow \mathbb{R}_+$  medible,

$$\mathbb{E}_x \left[ \int_0^{T_0} h(I_t, B_t) dt \right] = 2 \int_0^x \int_y^\infty h(y, z) dz dy.$$

Demostración: Notemos que

$$\begin{aligned} \mathbb{E}_x \left[ \int_0^{T_0} g(t, I_t, B_t) dt \right] &= \mathbb{E} \left[ \int_0^{T_x} g(t, I_t + x, B_t + x) dt \right] \\ &= \mathbb{E} \left[ \int_0^{T_x} g(t, x - S_t, x - B_t) dt \right] \\ &= \int_0^\infty \mathbb{E} [\mathbb{1}_{\{S_t < x\}} g(t, x - S_t, x - B_t)] dt \\ &= \int_0^\infty \int_0^x \int_{-\infty}^u g(t, x - u, x - v) \gamma_t(u, v) dv du dt \\ &= 2 \int_0^\infty \int_0^x \int_{-\infty}^u g(t, x - u, x - v) q_{2u-v}(t) dv du dt, \end{aligned}$$

con segunda igualdad dándose por Fubini y  $\{S_t < x\} = \{T_x > t\}$  bajo  $\mathbb{P}$  y la última igualdad por la identidad  $\gamma_t(u, v) = 2q_{2u-v}(t)$ . La primera igualdad del enunciado se obtiene al considerar el cambio de variable  $y = x - u$  y  $z = x - v$ . La segunda igualdad se debe a que  $t \mapsto q_{x+z-2y}(t)$  es una función de densidad. ■

Demostración de la proposición 2: Procedemos por inducción sobre  $p$ , notando que el caso  $p = 1$  se encuentra contenido en la parte (i) de la proposición 4.2. En esta demostración pondremos  $m_i = m(t_i, t_{i+1})$  y  $e_i = e(t_i)$ . Para el paso inductivo usamos la propiedad de Markov bajo  $n_+$ , la propiedad (ii) de la proposición 4.2, para obtener la primera igualdad en

$$\begin{aligned} n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \zeta\}} F(m_1, \dots, m_{p-1}, e_1, \dots, e_p) dt_1 \cdots dt_p \right) \\ = n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \zeta\}} dt_1 \cdots dt_{p-1} \mathbb{E}_{e_{p-1}} \left[ \int_0^{T_0} F(m_1, \dots, m_{p-2}, I_t, e_1, \dots, e_{p-1}, B_t) dt \right] \right) \\ = 2n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \zeta\}} dt_1 \cdots dt_{p-1} \int_0^{e_{p-1}} \int_{\alpha_{p-1}}^\infty F(m_1, \dots, m_{p-2}, \alpha_{p-1}, e_1, \dots, e_{p-1}, \beta_p) d\beta_p d\alpha_{p-1} \right), \end{aligned}$$

mientras que la segunda igualdad se debe al lema 3. El resultado se obtiene al usar la hipótesis de inducción. En específico, si definimos

$$\begin{aligned} G(\alpha_1, \dots, \alpha_{p-2}, \beta_1, \dots, \beta_{p-1}) \\ = \int_0^{\beta_{p-1}} \int_{\alpha_{p-1}}^\infty F(\alpha_1, \dots, \alpha_{p-1}, \beta_1, \dots, \beta_p) d\beta_p d\alpha_{p-1} \\ = \int_0^\infty \int_0^\infty \mathbb{1}_{\{\alpha_{p-1} \leq \beta_{p-1} \wedge \beta_p\}} F(\alpha_1, \dots, \alpha_{p-1}, \beta_1, \dots, \beta_p) d\beta_p d\alpha_{p-1}, \end{aligned}$$

aplicaremos la hipótesis de inducción sobre  $G$ , pues por lo derivado previamente tendremos que

$$\begin{aligned} n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \zeta\}} F(m_1, \dots, m_{p-1}, e_1, \dots, e_p) dt_1 \cdots dt_p \right) \\ = 2n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \zeta\}} G(m_1, \dots, m_{p-2}, e_1, \dots, e_{p-1}) \right. \\ \left. = 2^{p-1} \int_{\mathbb{R}_+^{2p-3}} \left( \prod_{i=1}^{p-2} \mathbb{1}_{\{\alpha_i \leq \beta_i \wedge \beta_{i+1}\}} \right) G(\alpha_1, \dots, \alpha_{p-2}, \beta_1, \dots, \beta_{p-1}) d\alpha_1 \cdots d\alpha_{p-2} d\beta_1 \cdots d\beta_{p-1}, \right. \end{aligned}$$

de donde concluimos gracias a la definición de  $G$ . ■

Demostración del teorema 1: Sea  $\Gamma_p$  la función medible de  $\mathbb{R}_+^{2p-1}$  en  $\Theta_{(p)}$  tal que

$$\theta(e; t_1, \dots, t_p) = \Gamma_p(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p)).$$

La existencia de esta función se sigue de la construcción inductiva del árbol marcado  $\theta(e; t_1, \dots, t_p)$ .

Sea  $\Delta_p$  la medida  $\mathbb{R}_+^{2p-1}$  definida por

$$\Delta_p(d\alpha_1 \cdots d\alpha_{p-1} d\beta_1 \cdots d\beta_p) = \prod_{i=1}^{p-1} \mathbb{1}_{[0, \beta_i \wedge \beta_{i+1}]}(\alpha_i) d\alpha_1 \cdots d\alpha_{p-1} d\beta_1 \cdots d\beta_p.$$

Debido a la proposición 2 únicamente habrá que verificar que  $\Lambda_p = (\Gamma_p)_* \Delta_p$ , es decir que  $\Lambda_p$  es la medida imagen de  $\Delta_p$  bajo  $\Gamma_p$ . Esta igualdad es clara para  $p = 1$ . Ciertamente, en tal caso, para una función  $F : \Theta_{(1)} \rightarrow \mathbb{R}_+$  medible,

$$\begin{aligned} n_+ \left( \int_{\{0 \leq t \leq \zeta\}} F(\theta(e; t)) dt \right) &= n_+ \left( \int_{\{0 \leq t \leq \zeta\}} F \circ \Gamma_1(e(t)) dt \right) \\ &= \int_{\mathbb{R}_+} F \circ \Gamma_1(t) dt \\ &= \int_{\mathbb{R}_+} F(\{\emptyset\}, t) dt = \int_{\Theta_{(1)}} F(\theta) d\Lambda_1(\theta). \end{aligned}$$

Procederemos con el resto de la prueba mediante inducción fuerte.

Sea pues  $p \geq 2$  y supongamos que el resultado es válido para todo  $j \in [p-1]$ . Para cada  $j \in [p-1]$  sea  $H_j$  el subconjunto de  $\mathbb{R}_+^{2p-1}$  definido como

$$H_j = \{(\alpha_1, \dots, \alpha_{p-1}, \beta_1, \dots, \beta_p) : \alpha_i > \alpha_j \text{ para toda } i \neq j\}.$$

Salvo un conjunto nulo bajo la medida de Lebesgue, los conjuntos  $H_j$  particionan el espacio  $\mathbb{R}_+^{2p-1}$ , de modo que

$$d\Delta_p(u) = \sum_{j=1}^{p-1} \mathbb{1}_{H_j}(u) d\Delta_p(u).$$

Ahora, se tiene que  $\mathbb{1}_{H_j}(u) d\Delta_p(u)$  es la imagen de la medida

$$\Upsilon_j = \Delta_j(d\alpha'_1 \cdots d\beta'_j) \otimes \mathbb{1}_{(0, \infty)}(h) dh \otimes \Delta_{p-j}(d\alpha''_1 \cdots d\beta''_{p-j})$$

bajo el mapeo  $\Phi : (\alpha'_1, \dots, \beta'_j, h, \alpha''_1, \dots, \beta''_{p-j}) \mapsto (\alpha_1, \dots, \beta_p)$  dado por

$$\begin{aligned} \alpha_i &= \begin{cases} h & \text{si } i = j, \\ \alpha'_i + h & \text{si } 1 \leq i \leq j-1, \\ \alpha''_{i-j} + h & \text{si } j+1 \leq i \leq p-1, \end{cases} \\ \beta_i &= \begin{cases} \beta'_i + h & \text{si } 1 \leq i \leq j, \\ \beta''_{i-j} + h & \text{si } j+1 \leq i \leq p. \end{cases} \end{aligned}$$

En vista de la construcción inductiva del árbol  $\theta(e; t_1, \dots, t_p)$  es claro que casi dondequiera con respecto a la medida  $\Upsilon_j$ , se cumple

$$\Gamma_p \circ \Phi(\alpha'_1, \dots, \beta'_j, h, \alpha''_1, \dots, \beta''_{p-j}) = \Gamma_j(\alpha'_1, \dots, \beta'_j) *_{\hbar} \Gamma_{p-j}(\alpha''_1, \dots, \beta''_{p-j}),$$

donde para  $\theta \in \Theta_{(j)}$  y  $\theta' \in \Theta_{(p-j)}$  se define a  $\theta *_{\hbar} \theta'$  como el árbol que se obtiene de concatenar los esqueletos discretos de  $\theta$  y  $\theta'$  y asignarle a  $\emptyset$  la marca  $h$ .

Entonces, por la hipótesis de inducción tendremos que para una función medible  $F$  sobre  $\Theta_{(p)}$ ,

$$\begin{aligned} & \int_{\mathbb{R}_+^{2p-1}} F \circ \Gamma_p(u) \mathbb{1}_{H_j}(u) d\Delta_p(u) \\ &= \int_0^\infty \int_{\mathbb{R}_+^{2j-1}} \int_{\mathbb{R}_+^{2(p-j)-1}} F \circ \Gamma_p \circ \Phi(u', h, u'') d\Delta_{p-j}(u'') d\Delta_j(u) dh \\ &= \int_0^\infty \int_{\mathbb{R}_+^{2j-1}} \int_{\mathbb{R}_+^{2(p-j)-1}} F(\Gamma_j(u') *_{\hbar} \Gamma_{p-j}(u'')) d\Delta_{p-j}(u'') d\Delta_j(u) dh \\ &= \int_0^\infty \int F(\theta) d(\Lambda_j *_{\hbar} \Lambda_{p-j})(\theta) dh, \end{aligned}$$

donde  $\Lambda_j *_{\hbar} \Lambda_{p-j}$  denota la imagen de la medida  $d\Lambda_j(\theta) d\Lambda_{p-j}(\theta')$  bajo la función  $(\theta, \theta') \mapsto \theta *_{\hbar} \theta'$ . Para completar la demostración resta notar que se tiene la igualdad

$$\Lambda_p = \sum_{j=1}^{p-1} \int_0^\infty (\Lambda_j *_{\hbar} \Lambda_{p-j}) dh.$$

Ella se cumple en vista de que si para  $\theta = (\tau, \{h_v : v \in \tau\}) \in \Theta_{(p)}^{\text{bin}}$ , si definimos  $\theta_1 = (\tau_1, \{h_v : v \in \tau_1\})$  mediante  $\tau_1 = \{v : 1v \in \tau\}$  y  $h_v = h_{1v}$  para  $1v \in \tau$ , y  $\theta_2$  de forma similar, entonces  $\theta = \theta_1 *_{\hbar} \theta_2$  y en particular omitiendo la etiqueta  $\emptyset$  en  $h_\emptyset$ ,

$$F(\theta) = \sum_{j=1}^{p-1} F(\theta_1 *_{\hbar} \theta_2) \mathbb{1}_{\{\theta_1 \in \Theta_{(j)}^{\text{bin}}\}} \mathbb{1}_{\{\theta_2 \in \Theta_{(p-j)}^{\text{bin}}\}}.$$

Así, usando que bajo  $\Lambda_p$  la etiqueta  $h$  es independiente del resto y se pone acorde a la medida de Lebesgue,

$$\begin{aligned} \int_{\Theta_{(p)}^{\text{bin}}} F(\theta) d\Lambda_p(\theta) &= \sum_{j=1}^{p-1} \int_0^\infty \int_{\Theta_{(j)}^{\text{bin}}} \int_{\Theta_{(p-j)}^{\text{bin}}} F(\theta_1 *_{\hbar} \theta_2) d\Lambda_{p-j}(\theta_2) d\Lambda_j(\theta_1) dh \\ &= \sum_{j=1}^{p-1} \int_0^\infty \int F(\theta) d(\Lambda_j *_{\hbar} \Lambda_{p-j})(\theta) dh. \end{aligned}$$

#### 4. Observación

Que en la demostración del teorema 1 se obtengan árboles binarios se sigue de que, con probabilidad uno, los mínimos locales del movimiento browniano son distintos. Formalmente, si  $0 < t_1 < \dots < t_p$  son dados y  $m_g(t_i, t_{i+1})$  son todos distintos, entonces el árbol  $\theta(g; t_1, \dots, t_p)$  resulta ser binario, situación que ocurre en el caso de una excursión browniana positiva.

## 5.6. Distribuciones finito-dimensionales del CRT

En la presente sección se realizará el cálculo de la ley de  $\theta(e; t_1, \dots, t_p)$ , donde recordamos que  $e$  es la excursión browniana normalizada. Esto es, el objetivo es determinar la ley del subárbol generado por  $p$  vértices escogidos de manera uniforme en el CRT. Veremos que para cada  $p \in \mathbb{N}$  obtendremos una medida de probabilidad sobre  $\Theta_{(p)}^{\text{bin}}$ , en lugar de la medida  $\Lambda_p$ , que llamaremos la distribución  $p$ -marginal del CRT.

Para realizar la construcción introduciremos una conexión entre la medida de Itô y la excursión browniana normalizada. De forma intuitiva, la ley de la excursión browniana normalizada es  $n(de \mid \zeta(e) = 1)$ . Debido a la descripción de Itô de la medida de excursiones positivas, ver por ejemplo el teorema XII.4.2 de Revuz y Yor (1999), y la propiedad de reescalamiento del movimiento browniano, sabemos que existe una única colección de medidas de probabilidad  $\{n_{(s)} : s > 0\}$  sobre el conjunto de excursiones  $\mathcal{E}$  tal que se cumplen:

- (i) Para cada  $s > 0$ ,  $n_{(s)}(\zeta = s) = 1$ .
- (ii) Dados  $\lambda > 0$  y  $s > 0$ , la ley de  $e_\lambda(t) = \sqrt{\lambda}e(t/\lambda)$  bajo  $dn_{(s)}(e)$  es  $n_{(\lambda s)}$ .
- (iii) Para cualquier subconjunto  $A \subset \mathcal{E}$  de Borel,

$$n_+(A) = \frac{1}{2\sqrt{2\pi}} \int_0^\infty \frac{1}{s^{3/2}} n_{(s)}(A) ds.$$

Resulta que la ley de la excursión browniana normalizada  $e$  es  $n_{(1)}$ .

El primera objetivo será extender el resultado que se presentó en el teorema 5.1, considerando el par  $(\theta(e; t_1, \dots, t_p), \zeta)$  en lugar de  $\theta(e; t_1, \dots, t_p)$ . Para un árbol marcado  $\theta$  definimos su longitud  $L(\theta)$  mediante

$$L(\theta) = \sum_{v \in \tau} h_v.$$

### 1. Proposición

La ley del par  $(\theta(e; t_1, \dots, t_p), \zeta)$  bajo la medida

$$dn_+(e) \mathbb{1}_{\{0 \leq t_1 \leq \dots \leq t_p \leq \zeta(e)\}} dt_1 \cdots dt_p$$

es

$$2^{p-1} d\Lambda_p(\theta) q_{2L(\theta)}(s) ds.$$

Demostración: Usaremos la misma notación que en la demostración del teorema 5.1. El paso clave de la demostración será verificar que para cualquier función medible  $F : \mathbb{R}_+^{3p} \rightarrow \mathbb{R}_+$  se cumple

$$\begin{aligned} & n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \zeta\}} dt_1 \cdots dt_p \right. \\ & \quad \left. f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p), t_1, t_2 - t_1, \dots, \zeta - t_p) \right) \\ (1) \quad & = 2^{p-1} \int_{\mathbb{R}_+^{2p-1}} \Delta_p(d\alpha_1 \cdots d\alpha_{p-1} d\beta_1 \cdots d\beta_p) \int_{\mathbb{R}_+^{p+1}} ds_1 \cdots ds_{p+1} q_{\beta_1}(s_1) \\ & \quad \times q_{\beta_1+\beta_2-2\alpha_1}(s_2) \cdots q_{\beta_{p-1}+\beta_p-2\alpha_{p-1}}(s_p) q_{\beta_p}(s_{p+1}) \\ & \quad \times f(\alpha_1, \dots, \alpha_{p-1}, \beta_1, \dots, \beta_p, s_1, \dots, s_{p+1}). \end{aligned}$$

En efecto, de demostrar la validez de (1), podemos usar la igualdad

$$2L(\Gamma_p(\alpha_1, \dots, \alpha_{p-1}, \beta_1, \dots, \beta_p)) = \beta_1 + \sum_{i=1}^{p-1} (\beta_i + \beta_{i-1} - 2\alpha_i) + \beta_p$$

junto a la identidad  $q_x * q_y = q_{x+y}$  para obtener

$$\begin{aligned} n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \zeta\}} f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p), \zeta) dt_1 \cdots dt_p \right) \\ = 2^{p-1} \int_{\mathbb{R}_+^{2p-1}} \int_0^\infty q_{2L(\Gamma_p(u))}(t) f(u, t) dt d\Delta_p(u) \end{aligned}$$

El enunciado, similar al teorema 5.1, es entonces una consecuencia de esta última igualdad y la identidad  $\Lambda_p = (\Gamma_p)_* \Delta_p$ .

Demostremos (1). Para el caso  $p = 1$  usaremos la propiedad de Markov bajo  $n_+$ , la definición de  $q_x$  y la parte (i) de la proposición 4.2 para obtener

$$\begin{aligned} \int_{\mathcal{E}} \int_0^\zeta f(e(t), t, \zeta - t) dt dn_+(e) \\ = \int_{\mathcal{E}} \int_0^\infty \mathbb{1}_{\{e(t) > 0\}} \mathbb{E}_{e(t)}[f(e(t), t, T_0)] dt dn_+(e) \\ = \int_{\mathcal{E}} \int_0^\infty \int_0^\infty \mathbb{1}_{\{e(t) > 0\}} q_{e(t)}(s) f(e(t), t, s) ds dt dn_+(e) \\ = \int_0^\infty \int_0^\infty \int_0^\infty q_x(t) q_x(s) f(x, t, s) ds dt dx, \end{aligned}$$

siendo esto la identidad deseada.

Consideremos ahora el caso  $p \geq 2$ . Aplicaremos, de manera sucesiva, la propiedad de Markov bajo  $n_+$  en los tiempos  $t_p$  y  $t_{p-1}$ , y finalmente aplicaremos el lema 5.3 para deducir

$$\begin{aligned} n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma\}} dt_1 \cdots dt_p \right. \\ \left. f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), \dots, e(t_1), \dots, e(t_p), t_1, t_2 - t_1, \dots, \zeta - t_p) \right) \\ = n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \sigma\}} dt_1 \cdots dt_{p-1} \int_{t_{p-1}}^\zeta dt_p \mathbb{E}_{e(t_p)} \left[ \right. \right. \\ \left. \left. f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), \dots, e(t_1), \dots, e(t_p), t_1, \dots, t_p - t_{p-1}, T_0) \right] \right) \\ = n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \zeta\}} dt_1 \cdots dt_{p-1} \int_{t_{p-1}}^\zeta dt_p \int_0^\infty ds q_{e(t)}(s) \right. \\ \left. f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), \dots, e(t_1), \dots, e(t_p), t_1, \dots, t_p - t_{p-1}, s) \right) \\ = n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \zeta\}} dt_1 \cdots dt_{p-1} \mathbb{E}_{e(t_{p-1})} \left[ \int_0^{T_0} dt \int_0^\infty ds q_{B_t}(s) \right. \right. \\ \left. \left. f(m(t_1, t_2), \dots, m(t_{p-2}, t_{p-1}), I_t, \dots, e(t_1), \dots, e(t_{p-1}), B_t, t_1, \dots, t_{p-1} - t_{p-2}, t, s) \right] \right) \\ = 2n_+ \left( \int_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \zeta\}} dt_1 \cdots dt_{p-1} \int_0^{e(t_{p-1})} dy \int_y^\infty dz \int_0^\infty dt q_{e(t_{p-1})+z-2y}(t) \int_0^\infty ds q_z(s) \right. \\ \left. f(m(t_1, t_2), \dots, m(t_{p-2}, t_{p-1}), y, e(t_1), \dots, e(t_{p-1}), z, t_1, \dots, t_{p-1} - t_{p-2}, t, s) \right). \end{aligned}$$

Al aplicar la hipótesis de inducción obtenemos (1), con lo que concluimos la demostración. ■

Gracias al resultado podemos enunciar y demostrar el resultado principal de la sección.

## 2. Teorema

La ley del árbol  $\theta(e; t_1, \dots, t_p)$  bajo la medida de probabilidad

$$p! \mathbb{1}_{\{0 \leq t_1 \leq \dots \leq t_p \leq 1\}} dt_1 \cdots dt_p dn_{(1)}(e)$$

es

$$p! 2^{p+1} L(\theta) \exp\{-2L(\theta)^2\} d\Lambda_p(\theta).$$

Demostración: Sea  $F$  una función no negativa, continua y acotada sobre  $\Theta_{(p)}$  y sea  $h$  una función acotada, no negativa y medible sobre  $\mathbb{R}_+$ . En vista de la proposición 1 tenemos la primera igualdad en

$$\begin{aligned} & \int_{\mathcal{E}} h(\zeta) \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \zeta\}} F(\theta(e; t_1, \dots, t_p)) dt_1 \cdots dt_p dn_+(e) \\ &= 2^{p-1} \int_0^\infty h(s) \int_{\Theta_{(p)}^{\text{bin}}} q_{2L(\theta)}(s) F(\theta) d\Lambda_p(\theta) ds \\ &= \frac{2^p}{\sqrt{2\pi}} \int_0^\infty \frac{1}{s^{3/2}} h(s) \int_{\Theta_{(p)}^{\text{bin}}} F(\theta) L(\theta) \exp\left\{-\frac{2L(\theta)^2}{s}\right\} d\Lambda_p(\theta) ds, \end{aligned}$$

con la segunda igualdad dándose por la definición de  $q_{2L(\theta)}(s)$ . Por otra parte, al usar la desintegración de la medida  $n_+$  dada por  $\{n_{(s)} : s > 0\}$ ,

$$\begin{aligned} & \int_{\mathcal{E}} h(\zeta) \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \zeta\}} F(\theta(e; t_1, \dots, t_p)) dt_1 \cdots dt_p dn_+(e) \\ &= \frac{1}{2\sqrt{2\pi}} \int_0^\infty \frac{1}{s^{3/2}} h(s) \int_{\mathcal{E}} \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq s\}} F(\theta(e; t_1, \dots, t_p)) dt_1 \cdots dt_p dn_{(s)}(e) ds. \end{aligned}$$

Comparando ambas identidades obtenemos que, casi dondequiera con respecto a  $s > 0$ ,

$$\begin{aligned} & 2^{p+1} \int_{\Theta_{(p)}^{\text{bin}}} F(\theta) L(\theta) \exp\left\{-\frac{2L(\theta)^2}{s}\right\} d\Lambda_p(\theta) \\ &= \int_{\mathcal{E}} \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq s\}} F(\theta(e; t_1, \dots, t_p)) dt_1 \cdots dt_p dn_{(s)}(e). \end{aligned}$$

Debido a la propiedad de reescalamiento de  $n_{(s)}$  se sigue que ambos lados de la igualdad son continuas como funciones de  $s$ , lo cual nos permite asegurar que son iguales para  $s = 1$ , situación que nos permite concluir. ■

Para concluir con la sección, y el capítulo, realizaremos algunas observaciones finales. Lo primero será notar que si escogemos  $t_1, \dots, t_p \in [0, 1]$  y consideramos el reacomodo creciente  $t'_1 \leq t'_2 \leq \dots \leq t'_p$ , entonces podemos definir  $\theta(e; t_1, \dots, t_p) = \theta(e; t'_1, \dots, t'_p)$ . Si en adición se tienen en cuenta del orden inicial, es posible definir el árbol etiquetado  $\tilde{\theta}(e; t_1, \dots, t_p)$  como el árbol  $\theta(e; t_1, \dots, t_p)$  donde la hoja correspondiente al tiempo  $t_i$  recibe la etiqueta  $i$ . De esta manera, el teorema 2 implica que la ley del árbol  $\tilde{\theta}(e; t_1, \dots, t_p)$ , bajo la medida de probabilidad

$$\mathbb{1}_{[0,1]^p}(t_1, \dots, t_p) dt_1 \cdots dt_p dn_{(1)}(e),$$

tiene densidad

$$2^{p+1}L(\theta) \exp\{-2L(\theta)^2\}$$

con respecto a la medida uniforme sobre  $\tilde{\Theta}_{(p)}^{\text{bin}}$ , el conjunto de árboles binarios, etiquetados y marcados con  $p$  hojas.

Ahora, dado un árbol en  $\tilde{\Theta}_{(p)}^{\text{bin}}$ , podemos “olvidar” el orden. Ponemos  $\hat{\theta}(e; t_1, \dots, t_p)$  como el árbol  $\tilde{\theta}(e; t_1, \dots, t_p)$  sin la estructura de orden. Como para cada árbol binario etiquetado con  $p$  hojas hay  $2^{p-1}$  órdenes posibles, deducimos que la ley del árbol  $\hat{\theta}(e; t_1, \dots, t_p)$  tiene densidad

$$2^{2p}L(\theta) \exp\{-2L(\theta)^2\}$$

con respecto a la medida uniforme sobre  $\hat{\Theta}_{(p)}^{\text{bin}}$ , el conjunto de árboles binarios, etiquetados y marcados, sin orden, con  $p$  hojas.

De lo anterior, y realizando cuentas similares, se verifica que si se reemplaza la excursión  $e$  por  $2e$ , lo cual coincide con la normalización del CRT dada por Aldous (1991a, 1991b, 1993), entonces la ley del árbol  $\hat{\theta}(2e; t_1, \dots, t_p)$  tiene densidad

$$(2) \quad L(\theta) \exp\left\{-\frac{L(\theta)^2}{2}\right\}$$

con respecto a la medida uniforme sobre  $\hat{\Theta}_{(p)}^{\text{bin}}$ . De hecho, en la forma (2) es posible reconocer la densidad de las distribuciones finito-dimensionales propuestas por Aldous, ver por ejemplo el lema 21 del artículo «The Continuum Random Tree III». En efecto, al realizar la descripción explícita vemos que el árbol tiene un esqueleto discreto  $\hat{\tau}(2e; t_1, \dots, t_p)$  cuya distribución es uniforme sobre el conjunto de árboles binarios con  $p$  hojas etiquetadas, el cual tiene  $b_p$  elementos, donde

$$b_p = p!2^{-p+1}c_{p-1} = \prod_{i=1}^p (2i-3).$$

Luego, condicional en el esqueleto, las marcas  $h_v$  tienen distribución con la densidad

$$(3) \quad b_p \sum h_v \exp\left\{-\frac{1}{2}\left(\sum h_v\right)^2\right\}.$$

Que (3) en efecto define una densidad se puede verificar al notar que

$$\begin{aligned} \int_{\mathbb{R}_+^{2p-1}} \sum_{v=1}^{2p-1} h_v \exp\left\{-\frac{1}{2}\left(\sum_{v=1}^{2p-1} h_v\right)^2\right\} \prod_{v=1}^{2p-1} dh_v &= \int_{\mathbb{R}_+^{2p-2}} \exp\left\{-\frac{1}{2}\left(\sum_{v=1}^{2p-2} h_v\right)^2\right\} \prod_{v=1}^{2p-2} dh_v \\ &= \frac{1}{\Gamma(2p-2)} \int_{\mathbb{R}_+} r^{2p-3} \exp\left\{-\frac{r^2}{2}\right\} dr \\ &= \frac{2^{p-2}(p-2)!}{(2p-3)!} = \frac{1}{b_p}. \end{aligned}$$

De esta manera concluimos con el estudio de las distribuciones finito-dimensionales del CRT.

## Parte II.

# Genealogía de una muestra

En esta parte se realiza un estudio detallado de artículo «The Coalescent Structure of Continuous-Time Galton–Watson Trees» de Harris et al. (2020), en el cual se presentan resultados sobre la genealogía de  $k$  individuos dentro de una población dada por un árbol de Galton–Watson. Para obtener los resultados, los autores desarrollan la técnica de  $k$  espinas, la cual extiende la técnica de 1 espina usada por Lyons et al. (1995) en su artículo «Conceptual Proofs of  $L \log L$  Criteria for Mean Behavior of Branching Processes».

# Introducción

En la sección 6.1 introducimos el árbol de Galton–Watson con el cual trabajaremos a lo largo de esta parte, así como estableceremos la notación que se usará en la misma. Adicionalmente, en la sección 6.2 se incluyen algunos resultados conocidos de procesos de Galton–Watson que se usarán en capítulos posteriores.

## 6.1. El modelo base

El modelo básico que usaremos en esta parte del texto es aquel de un árbol de Galton–Watson con tiempo continuo. Considerando los elementos desarrollados en el capítulo 1, un árbol de Galton–Watson en tiempo continuo puede entenderse como un árbol de Galton–Watson marcado, donde las marcas corresponden a los tiempos de vida de los individuos.

De manera más formal, bajo una medida de probabilidad  $\mathbb{P}$ , comenzamos con un individuo, la raíz, que tiene etiqueta  $\emptyset$ . Este individuo tiene un tiempo de vida  $\tau_\emptyset$  con distribución exponencial de parámetro  $r > 0$ . En el tiempo  $\tau_\emptyset$  el individuo se reproduce de acuerdo a una variable aleatoria  $L$  con soporte en  $\mathbb{N}_0$  e inmediatamente muere. A los descendientes del individuo  $\emptyset$  se les asignan las etiquetas,  $1, 2, \dots, L_\emptyset$ , donde  $L_\emptyset$  es una copia independiente de  $L$ . Ellos repiten, de manera independiente, el comportamiento del individuo original; a saber, cada individuo  $u$  tiene un tiempo de vida con distribución exponencial de parámetro  $r$ , da a luz a los individuos  $u1, \dots, uL_u$  y muere. Ponemos  $p_j = \mathbb{P}(L = j)$  para toda  $j \in \mathbb{N}_0$  y  $m = \sum_{j \geq 1} j p_j$ . En adición, debido a que se usarán distintas medidas de probabilidad, escribiremos  $\mathbb{P}[\cdot]$  en lugar de  $\mathbb{E}[\cdot]$  para denotar la esperanza bajo la medida de probabilidad  $\mathbb{P}$ .

Para un tiempo  $T > 0$ , denotamos por  $\mathcal{N}_T$  al conjunto de individuos vivos al tiempo  $T$  y por  $N_T = \#\mathcal{N}_T$  al número de individuos vivos al tiempo  $T$ . Además, para un individuo  $u \in \mathcal{N}_T$  definimos a  $\tau_u$  como el tiempo de su muerte y  $\tau_u(T) = \tau_u \wedge T$ . Mantenemos la notación establecida en el capítulo 1, en la cual  $v \leq u$  si  $v$  es ancestro de  $u$  y  $v < u$  si  $v$  es ancestro *estricto* de  $u$ . También añadimos un cementerio  $\dagger$ , el cual no está vivo y por ende no puede ser elemento de  $\mathcal{N}_T$ .

Dado un individuo  $u \in \mathcal{N}_t$  y  $s \leq t$ , definimos por  $u(s)$  al único ancestro de  $u$  vivo al tiempo  $s$ . Mientras tanto, dados dos individuos  $u, v \in \mathcal{N}_T$  consideramos a  $\sigma(u, v)$  como el último tiempo en el cual tienen un ancestro en común, es decir

$$\sigma(u, v) := \sup\{t \geq 0 : u(t) = v(t)\}.$$

Ahora, dados  $k \in \mathbb{N}$  y un tiempo  $T > 0$ , en el evento  $\{N_T \geq k\}$  escogemos  $k$  individuos  $U_T^1, \dots, U_T^k$  de manera uniforme y sin reemplazo del conjunto  $\mathcal{N}_T$ . Esto nos permite definir para cada  $t \in [0, T]$ , una partición de  $[k]$  considerando que  $i$  y  $j$  se encuentran relacionados,  $i \sim j$ , precisamente cuando  $U_T^i(t) = U_T^j(t)$ ; es decir cuando  $U_T^i$  y  $U_T^j$  tienen un ancestro en común al tiempo  $t$ , lo cual es equivalente

a que  $\sigma(U_T^i, U_T^j) > t$ . A tal partición la denotamos por  $\mathcal{P}_t^k(T)$  y ordenaremos a sus clases acorde a los menores elementos de cada una de ellas.

Notemos que la colección  $\{\mathcal{P}_t^k(T) : t \in [0, T]\}$  nos da la siguiente información:

- Por una parte contiene la topología del árbol hasta el tiempo  $T$ . En efecto, dada una colección de bloques, nos dice qué bloque se separa primero y cómo se ven los bloques que se crean en tal separación.
- Por otra parte nos dice los tiempos en los cuales ocurren las separaciones.

El separar ambos aspectos se debe a que la información topológica de los modelos tratados por Harris et al. (2020) es asintóticamente parecida a la del coalescente de Kingman, mientras que los tiempos de separación son sensibles a los parámetros del modelo.

Por ende, con el fin de enfatizar la separación de la topología del árbol y los tiempos de separación, impondremos más notación. Definimos a  $v_t^k(T)$  como el número de bloques en  $\mathcal{P}_t^k(T)$ , o bien el número de ancestros de  $U_T^1, \dots, U_T^k$  en el tiempo  $t$ . Formalmente esto se expresa como

$$v_t^k(T) := \#\{u \in \mathcal{N}_t : u \leq U_T^i \text{ para algún } i \leq k\}.$$

La colección  $\{v_t^k(T) : t \in [0, T]\}$  nos da la información topológica, de modo que falta definir variables que nos den la información de los tiempos de separación. Para  $i \in [k-1]$  ponemos

$$S_i^k(T) := \inf\{t \geq 0 : v_t^k(T) > i\}.$$

A los tiempos  $S_1^k(T) \leq \dots \leq S_{k-1}^k(T)$  los llamaremos *tiempos de división*. Asociados a los tiempos de división, definimos los tiempos de división desordenados  $(\tilde{S}_1^k(T), \dots, \tilde{S}_{k-1}^k(T))$ , que es únicamente una permutación aleatoria y uniforme de  $(S_1^k(T), \dots, S_{k-1}^k(T))$ . Finalmente, para cada  $i \in \{0, \dots, k-1\}$  definimos la partición aleatoria  $\mathcal{P}_i^k(T) = \mathcal{P}_{S_i^k(T)}^k(T)$ , donde  $S_0^k(T) = 0$ , y definimos  $\mathcal{H} = \sigma(\mathcal{P}_0^k(T), \dots, \mathcal{P}_{k-1}^k(T))$ . La  $\sigma$ -álgebra  $\mathcal{H}$  contiene toda la información de la topología del árbol, pero la única información que contiene de los tiempos de división es si hay empates entre ellos.

## 6.2. Resultados auxiliares

En esta sección recopilamos algunos resultados conocidos sobre procesos de Galton–Watson. El primero de ellos concierne al caso crítico, donde  $m = 1$ .

### 1. Teorema

Bajo el supuesto  $m = 1$  y poniendo  $\sigma^2 := \text{Var}(L) = \mathbb{P}[L^2] - 1$  tenemos:

**Estimación de Kolmogorov**  $\mathbb{P}(N_T > 0) \rightarrow 2/\sigma^2$  conforme  $T \rightarrow \infty$ ;

**Ley límite de Yaglom** si  $\sigma < \infty$ , entonces la distribución condicional de  $N_T/T$  dado  $N_T > 0$  converge, conforme  $T \rightarrow \infty$ , a una ley exponencial de parámetro  $2/\sigma^2$ .

La demostración de este resultado, en el caso discreto, se puede consultar en la sección 12.4 del libro *Probability on Trees and Networks* de Lyons y Peres (2016) y en la demostración que presentan usan la idea de 1 espina.

**TODO:** Ir escribiendo resultados que se usen a lo largo de la parte.

# Espinas y cambios de medidas

# 7

Este capítulo es el más técnico de esta parte, pues aquí es donde se establecen las herramientas que se requerirán para demostrar resultados en los capítulos posteriores. De los resultados más relevantes que se exponen a lo largo del capítulo son las proposiciones 2.2 y 6.1. El primero de ellos proporciona una traducción entre preguntas relacionadas a individuos escogidos de manera aleatoria bajo  $\mathbb{P}$  y cálculos bajo una nueva medida  $\mathbb{Q}$ , mientras que el segundo es una versión de la fórmula de Campbell bajo  $\mathbb{Q}$ .

Claro, lo primero que se debe hacer es introducir la medida  $\mathbb{Q}$ , para lo cual describiremos la noción de *espina*, concepto que introduce información extra al árbol aleatorio asignando marcas especiales a algunos individuos.

## 7.1. La medida de $k$ -espinas

Comenzamos definiendo una nueva medida  $\mathbb{P}^k$  bajo la cual habrá  $k$  líneas ancestrales distinguidas. De manera concisa,  $\mathbb{P}^k$  es una extensión de  $\mathbb{P}$  en la cual algunos individuos tienen una marca especial que indica que ellos son parte de la espina. Más precisamente, bajo  $\mathbb{P}^k$  los individuos actúan como sigue:

- Se comienza con un individuo con  $k$  marcas especiales, denotadas por  $1, \dots, k$ .
- Cada una de las marcas especiales  $1, \dots, k$  se entiende como una línea ancestral distinguida o espina, y definimos a  $\xi_t^i$  como la etiqueta del individuo que tenga la marca especial  $i$  al tiempo  $t$ .
- Un individuo con  $j$  marcas especiales,  $b_1 < \dots < b_j$ , al tiempo  $t$ , se ramifica a tasa  $r$ , muriendo y siendo reemplazado por un número aleatorio de individuos con la misma ley que  $L$ , de manera independiente al resto del sistema.
- Dados que  $a$  individuos  $v_1, \dots, v_a$  nacen en el evento de ramificación del punto anterior, cada una de las  $j$  marcas especiales se asigna de manera uniforme, en independiente al resto, a uno de los  $a$  individuos. Así, para cada  $l \in [a]$  e  $i \in [j]$ , la probabilidad de que se le asigne la marca especial  $b_i$  al individuo  $v_l$  es  $1/a$ , independientemente del resto de las asignaciones.
- Si un individuo con  $j$  marcas,  $b_1 < \dots < b_j$ , muere y no tiene descendientes, entonces las  $j$  marcas especiales se le asignan al cementerio  $\dagger$ .

De la descripción anterior, es claro que el sistema de individuos tiene el mismo comportamiento bajo  $\mathbb{P}^k$  que bajo  $\mathbb{P}$ , solamente se añade la información de las marcas especiales que nos dicen las líneas ancestrales correspondientes a las  $k$  espinas. Definimos  $\xi_t := (\xi_t^1, \dots, \xi_t^k)$  a la colección de etiquetas asociadas a las marcas especiales.

En adición denotamos por  $n_t$  al número de espinas distintas, el número de individuos con marcas especiales, al tiempo  $t$  y, para  $i \geq 1$ ,

$$\psi_i := \inf\{t \geq 0 : n_t \notin [i]\},$$

con  $\psi_0 = 0$ . La variable  $\psi_i$  se puede entender como el tiempo de la  $i$ -ésima división de las espinas, aunque hay que tener cuidado con la nomenclatura, pues es posible que  $\psi_i = \psi_{i+1}$ . También ponemos  $\rho_t^i$  para referirnos al número de marcas especiales sobre la espina  $i$ .

Observemos que en cualquier tiempo  $t$ , el conjunto de individuos distintos sobre las espinas induce una partición  $\mathcal{Z}_t^k$  sobre  $[k]$ . En efecto, podemos considerar que  $i$  y  $j$  se encuentran en el mismo bloque de  $\mathcal{Z}_t^k$  si  $\xi_t^i = \xi_t^j$ . Entonces, si definimos  $Z_i^k = \mathcal{Z}_{\psi_i}^k$  para  $i \in \{0, \dots, k-1\}$ , donde  $Z_0^k = [k]$ , tendremos que  $\{Z_0^k, \dots, Z_{k-1}^k\}$  conforma una colección de particiones que describe la topología de las espinas, mientras que de los tiempos de división de las espinas únicamente contiene información sobre empates entre los tiempos de división. De manera similar a  $\mathcal{H}$ , definimos  $\mathcal{H}' := \sigma(Z_0^k, Z_1^k, \dots)$ .

Dada un individuo  $u \in \mathcal{N}_t$ , es claro que existe un último tiempo  $s \in [0, t]$  en el cual  $u(s)$  se encuentra sobre una espina. En caso de que este tiempo sea igual a  $\psi_i$  para algún  $i$ , diremos que  $u$  es un *individuo residual*. Si, en cambio, este tiempo es distinto de  $\psi_i$  para toda  $i$  y  $u$  no se encuentra sobre una espina, entonces diremos que es un *individuo ordinario*. Esto nos permite clasificar a todos los individuos como residuales, ordinarios o espinas (que se encuentran sobre una espina).

Antes de proceder a la siguiente sección, cabe destacar que  $\mathbb{P}^k$  no se define en la misma  $\sigma$ -álgebra que  $\mathbb{P}$ , sino que está definida en  $\mathcal{F}_\infty^k = \bigvee_{t \geq 0} \mathcal{F}_t^k$ , donde  $\mathcal{F}_t^k$  es la  $\sigma$ -álgebra que contiene la información del sistema y las  $k$  espinas hasta el tiempo  $t$ . Adicionalmente, para cada  $t \geq 0$  consideramos las  $\sigma$ -álgebras  $\mathcal{F}_t^0$ ,  $\tilde{\mathcal{G}}_t^k$  y  $\mathcal{G}_t^k$ . La  $\sigma$ -álgebra  $\mathcal{F}_t^0$  contiene la información del árbol de Galton–Watson,  $\tilde{\mathcal{G}}_t^k$  la información de las  $k$  espinas incluyendo los eventos de nacimiento sobre las  $k$  espinas (sin información adicional del resto del árbol) y  $\mathcal{G}_t^k$  la información de los eventos de división de las espinas incluyendo la distribución de las marcas sobre las espinas. Por definición  $\mathcal{G}_t^k$  no tiene información sobre el nacimiento de individuos ordinarios.

## 7.2. Cambio de medida

Habiendo introducido las  $k$  espinas, ahora el objetivo es introducir una nueva medida bajo la cual ellas se escojan de manera uniforme y sin reemplazo en el tiempo  $T$ . Esto nos permitirá representar individuos escogidos de manera uniforme bajo  $\mathbb{P}$  usando las espinas bajo esta nueva medida. Veremos que esta nueva medida puede ser completamente descrita hacia adelante en el tiempo.

En lo que resta del capítulo fijamos  $k \in \mathbb{N}$ . Además, hasta nuevo aviso, supondremos que  $\mathbb{P}[L^k] < \infty$ .

Para cualquier conjunto  $S$  y  $k \geq 1$  definimos a  $S^{(k)} \subset S^k$  como el conjunto de  $k$ -tuplas de  $S$  con todas las entradas distintas. Por otra parte, para  $n \geq 0$  definimos

$$n^{(k)} = \frac{n!}{(n-k)!} = n(n-1) \cdots (n-k+1) \quad \text{si } n \geq k,$$

y  $n^{(k)} = 0$  en otro caso. Con estas notaciones, es fácil ver que  $\#(S^{(k)}) = (\#S)^{(k)}$ .

Para enunciar el primer resultado del presente apartado introducimos, para  $t \geq 0$ , las variables aleatorias

$$(1) \quad g_{k,t} := \mathbb{1}_{\{\xi_t \in \mathcal{N}_t^{(k)}\}} \prod_{i=1}^k \prod_{v < \xi_t^i} L_v,$$

y

$$\zeta_{k,t} := \frac{g_{k,t}}{\mathbb{P}[N_t^{(k)}]}.$$

### 1. Lema

Para cada  $t \geq 0$ ,

$$\mathbb{P}^k[g_{k,t} \mid \mathcal{F}_t^0] = N_t^{(k)}.$$

En particular se cumple que  $\mathbb{P}^k[\zeta_{k,t}] = 1$ .

Demostración: Notemos que  $\mathbb{1}_{\{\xi_t \in \mathcal{N}_t^{(k)}\}} = \sum_{u \in \mathcal{N}_t^{(k)}} \mathbb{1}_{\{\xi_t = u\}}$ , de modo que por definición de  $g_{k,t}$ , linealidad y que dado  $u \in \mathcal{N}_t^{(k)}$ ,  $\prod_{i=1}^k \prod_{v < u_i} L_v$  es  $\mathcal{F}_t^0$ -medible,

$$\begin{aligned} \mathbb{P}^k[g_{k,t} \mid \mathcal{F}_t^0] &= \mathbb{P}^k \left[ \sum_{u \in \mathcal{N}_t^{(k)}} \mathbb{1}_{\{\xi_t = u\}} \prod_{i=1}^k \prod_{v < u_i} L_v \mid \mathcal{F}_t^0 \right] \\ &= \sum_{u \in \mathcal{N}_t^{(k)}} \left( \prod_{i=1}^k \prod_{v < u_i} L_v \right) \mathbb{P}^k(\xi_t = u \mid \mathcal{F}_t^0). \end{aligned}$$

Recordando que las marcas especiales actúan de manera independiente y en cada evento de ramificación escogen un individuo entre los descendientes de manera uniforme, se sigue que

$$\mathbb{P}^k(\xi_t = u \mid \mathcal{F}_t^0) = \prod_{i=1}^k \mathbb{P}^k(\xi_t^i = u_i \mid \mathcal{F}_t^0) = \prod_{i=1}^k \prod_{v < u_i} \frac{1}{L_v},$$

lo cual nos permite deducir

$$\mathbb{P}^k[g_{k,t} \mid \mathcal{F}_t^0] = \sum_{u \in \mathcal{N}_t^{(k)}} 1 = \#\mathcal{N}_t^{(k)} = N_t^{(k)}.$$

De esta manera obtenemos la primera parte del enunciado, de donde se sigue inmediatamente la segunda parte al tomar esperanza bajo  $\mathbb{P}^k$ . ■

Ahora fijamos  $T > 0$  y usamos  $\zeta_{k,T}$  para definir una nueva medida de probabilidad  $\mathbb{Q}^{k,T}$  mediante

$$(2) \quad \frac{d\mathbb{Q}^{k,T}}{d\mathbb{P}^k} \Big|_{\mathcal{F}_T^k} := \frac{N_T(N_T - 1) \cdots (N_T - k + 1)}{\mathbb{P}[N_T(N_T - 1) \cdots (N_T - k + 1)]} = \frac{N_T^{(k)}}{\mathbb{P}[N_T^{(k)}]} = \zeta_{k,T}.$$

Entonces, gracias al lema 1, se tiene que

$$(3) \quad \frac{d\mathbb{Q}^{k,T}}{d\mathbb{P}^k} \Big|_{\mathcal{F}_T^0} = \frac{N_T^{(k)}}{\mathbb{P}[N_T^{(k)}]} =: Z_{k,T}.$$

El siguiente resultado, cuya demostración se pospone hasta la sección 7.5, nos dice que la medida  $\mathbb{Q}^{k,T}$  es útil para traducir cuestiones sobre individuos escogidos de manera uniforme bajo  $\mathbb{P}$  en preguntas sobre las espigas bajo la medida  $\mathbb{Q}^{k,T}$ .

## 2. Proposición

Supongamos que  $f$  es una función medible de  $k$ -tuplas de partículas al tiempo  $T$ . Entonces

$$\begin{aligned} \mathbb{P} \left[ \frac{1}{N_T^{(k)}} \sum_{u \in \mathcal{N}_T^{(k)}} f(u) \mid N_T \geq k \right] \\ = \frac{\mathbb{P}[N_T^{(k)}]}{\mathbb{P}(N_T \geq k)(k-1)!} \int_0^\infty (e^z - 1)^{k-1} \mathbb{Q}^{k,T}[e^{-zN_T} f(\xi_T)] dz. \end{aligned}$$

## 7.3. Descripción de $\mathbb{Q}^{k,T}$

En este apartado damos una descripción de la medida  $\mathbb{Q}^{k,T}$  a través de una serie de resultados. En la demostración de las propiedades, será de utilidad el siguiente lema.

### 1. Lema

Supongamos que  $\mu$  y  $\nu$  son medidas de probabilidad sobre la  $\sigma$ -álgebra  $\mathcal{F}$  tales que  $\mu \ll \nu$  y que  $\mathcal{G} \subset \mathcal{F}$  es una sub- $\sigma$ -álgebra. Si

$$\frac{d\mu}{d\nu} \Big|_{\mathcal{F}} = Y \quad \text{y} \quad \frac{d\mu}{d\nu} \Big|_{\mathcal{G}} = Z,$$

entonces para cualquier variable aleatoria no negativa  $X$ ,

$$Z\mu[X \mid \mathcal{G}] = \nu[XY \mid \mathcal{G}] \quad \nu\text{-casi seguramente.}$$

Demostración: Para cualquier  $A \in \mathcal{G}$  se tiene, por definición de la derivada de Radon–Nikodym,

$$\nu[XY\mathbb{1}_A] = \mu[X\mathbb{1}_A] = \mu[\mu[X \mid \mathcal{G}]\mathbb{1}_A] = \nu[Z\mu[X \mid \mathcal{G}]\mathbb{1}_A].$$

El resultado se sigue en vista de que  $Z\mu[X \mid \mathcal{G}]$  es  $\mathcal{G}$ -medible. ■

El lema 1 junto a (2.2) y (2.3) nos dice que para cualquier variable aleatoria no negativa  $X$ , medible con respecto a  $\mathcal{F}_T^k$ , se cumple la igualdad

$$(1) \quad \mathbb{Q}^{k,T}[X \mid \mathcal{F}_T^0] = \frac{1}{Z_{k,T}} \mathbb{P}^k[X\zeta_{k,T} \mid \mathcal{F}_T^0]$$

en el evento  $\{Z_{k,T} > 0\}$ . En adición, como  $\zeta_{k,T}$  es medible con respecto a  $\tilde{\mathcal{G}}_T^k$ , en el evento  $\{\zeta_{k,T} > 0\}$  tendremos la igualdad

$$(2) \quad \mathbb{Q}^{k,T}[X \mid \tilde{\mathcal{G}}_T^k] = \frac{1}{\zeta_{k,T}} \mathbb{P}^k[X\zeta_{k,T} \mid \tilde{\mathcal{G}}_T^k] = \mathbb{P}^k[X \mid \tilde{\mathcal{G}}_T^k].$$

La expresión (2) nos dice que cualquier evento que sea independiente de  $\tilde{\mathcal{G}}_T^k$  tiene la misma probabilidad bajo  $\mathbb{Q}^{k,T}$  que bajo  $\mathbb{P}^k$ . En palabras llanas *los individuos no espinales tiene el mismo comportamiento bajo  $\mathbb{Q}^{k,T}$  que bajo  $\mathbb{P}^k$ .*

El primer resultado que presentamos afirma que  $\mathbb{Q}^{k,T}$  satisface una propiedad de Markov de ramificación que depende del tiempo, en el sentido de que los descendientes de cualquier individuo actúan de manera independiente al resto del árbol.

**2. Lema** (de simetría)

*Supongamos que  $u \in \mathcal{N}_t$  tiene  $j$  marcas especiales al tiempo  $t$ . Entonces, bajo  $Q^{k,T}$ , el subárbol generado por  $v$  después del tiempo  $v$  es independiente del resto del sistema y es descrito por la medida  $Q^{j,T-t}$ .*

Demostración: Fijemos  $0 < t < T$  y  $v \in \mathcal{N}_t$ . Sea  $\mathcal{H}_t^v$  la  $\sigma$ -álgebra generada por toda la información salvo aquella del subárbol generado por  $v$  tras el tiempo  $t$ . Entonces basta probar que para  $s \in (t, T]$  e  $i \geq 0$  se cumple, con probabilidad uno, que

$$Q^{k,T}(\tau_v > s, L_v = i \mid \mathcal{H}_t^v) = Q^{k,T-t}(\tau_\emptyset > s - t, L_\emptyset = i).$$

Si denotamos por  $I_v$  al conjunto de marcas especiales que se encuentran asignadas a  $v$  al tiempo  $t$ , y definimos

$$\tilde{g} = \mathbb{1}_{\{\xi_T^i \neq \xi_T^j, i, j \in I_v, i \neq j\}} \prod_{i \in I_v} \prod_{\xi_t^i \leq u < \xi_T^i} L_u$$

y

$$h = \mathbb{1}_{\{\xi_T^i \neq \xi_T^j, i, j \in I_v, i \neq j\}} \left( \prod_{i \in I_v} \prod_{u < \xi_T^i} L_u \right) \left( \prod_{i \in I_v} \prod_{u < \xi_T^i} L_u \right),$$

por definición  $h$  es  $\mathcal{H}_t^v$ -medible y también se cumple  $g_{k,T} = \tilde{g}h$ .

Así, debido al lema 1,  $Q^{k,T}$ -casi seguramente,

$$Q^{k,T}(\tau_v > s, L_v = i \mid \mathcal{H}_t^v) = \frac{1}{\mathbb{P}^k[\zeta_{k,T} \mid \mathcal{H}_t^v]} \mathbb{P}^k[\zeta_{k,T} \mathbb{1}_{\{\tau_v > s, L_v = i\}} \mid \mathcal{H}_t^v].$$

Luego, al usar la definición de  $\zeta_{t,K}$  podemos cancelar los factores  $\mathbb{P}^k[N_t^{(k)}]$  para obtener

$$Q^{k,T}(\tau_v > s, L_v = i \mid \mathcal{H}_t^v) = \frac{1}{\mathbb{P}^k[g_{k,T} \mid \mathcal{H}_t^v]} \mathbb{P}^k[g_{k,T} \mathbb{1}_{\{\tau_v > s, L_v = i\}} \mid \mathcal{H}_t^v],$$

y entonces, como  $g_{k,T} = \tilde{g}h$  y  $h$  es  $\mathcal{H}_t^v$ -medible deducimos

$$Q^{k,T}(\tau_v > s, L_v = i \mid \mathcal{H}_t^v) = \frac{1}{\mathbb{P}^k[\tilde{g} \mid \mathcal{H}_t^v]} \mathbb{P}^k[\tilde{g} \mathbb{1}_{\{\tau_v > s, L_v = i\}} \mid \mathcal{H}_t^v].$$

Observemos que la propiedad de Markov de la ramificación bajo  $\mathbb{P}^k$  nos dice que el subárbol generado por  $v$  tras el tiempo  $t$  es independiente del resto del sistema y en adición, en el evento en el cual  $v$  tiene asignadas  $j$  marcas especiales, la distribución del subárbol es aquella dada por  $\mathbb{P}^j$ . Puesto que  $\tilde{g}$  depende del subárbol generado por  $v$  tras el tiempo  $t$ , lo anterior implica que obtenemos las identidades casi seguras

$$\mathbb{P}^k[\tilde{g} \mid \mathcal{H}_t^v] = \mathbb{P}^j[g_{j,T-t}] \quad \text{y} \quad \mathbb{P}^k[\tilde{g} \mathbb{1}_{\{\tau_v > s, L_v = i\}} \mid \mathcal{H}_t^v] = \mathbb{P}^j[g_{j,T-t} \mathbb{1}_{\{\tau_\emptyset > s-t, L_\emptyset = i\}}],$$

con las cuales obtenemos

$$\begin{aligned} Q^{k,T}(\tau_v > s, L_v = i \mid \mathcal{H}_t^v) &= \frac{\mathbb{P}^j[g_{j,T-t} \mathbb{1}_{\{\tau_\emptyset > s-t, L_\emptyset = i\}}]}{\mathbb{P}^j[g_{j,T-t}]} \\ &= \mathbb{P}^j[\zeta_{j,T-t} \mathbb{1}_{\{\tau_\emptyset > s-t, L_\emptyset = i\}}] = Q^{j,T-t}(\tau_\emptyset > s - t, L_\emptyset = i), \end{aligned}$$

donde la segunda igualdad se da gracias al lema 2.1. ■

Como ya se mencionó, (2) implica que los individuos que no son espinas se comportan como bajo  $\mathbb{P}^k$ . El comportamiento de los individuos sobre la espina es más complicado de describir. No obstante, en virtud del lema de simetría, para entender los tiempos de división bajo  $Q^{k,T}$ , basta entender las distribuciones de  $\tau_\emptyset$  y de  $\psi_1$ . Este es el contenido del siguiente resultado.

### 3. Lema

Para cualesquiera  $t \in [0, T]$  y  $k \geq 0$  tenemos

$$\mathbb{Q}^{k,T}(\tau_\emptyset > t) = \frac{\mathbb{P}^k[N_{T-t}^{(k)}]}{\mathbb{P}^k[N_T^{(k)}]} e^{-rt},$$

$$\mathbb{Q}^{k,T}(\psi_1 > t) = \frac{\mathbb{P}^k[N_{T-t}^{(k)}]}{\mathbb{P}^k[N_T^{(k)}]} e^{(m-1)rt}$$

y

$$\mathbb{Q}^{k,T}(\tau_\emptyset > t \mid \psi_1 > t) = e^{-mrt}.$$

Previo a realizar la prueba, notemos que la tercera parte del lema 3 junto al lema de simetría, nos dice que dada  $\mathcal{G}_T^k$ , la información de los tiempos de división de las espinas, bajo  $\mathbb{Q}^{k,T}$  cada espina da a luz a individuos que no son espinas acorde a un proceso Poisson de tasa  $mr$ , de forma independiente al resto del sistema. En particular, cuando haya  $n$  espinas distintas, la tasa a la cual aparecen individuos que no son espinas es  $nmr$ . Dicho esto, demostremos el lema 3.

Demostración: Comenzamos notando que

$$\mathbb{Q}^{k,T}(\tau_\emptyset > t) = \mathbb{P}^k[\zeta_{k,T} \mathbb{1}_{\{\tau_\emptyset > t\}}] = \frac{1}{\mathbb{P}[N_T^{(k)}]} \mathbb{P}[g_{k,T} \mathbb{1}_{\{\tau_\emptyset > t\}}].$$

Entonces, como por la propiedad de Markov y el lema 2.1 se cumple

$$\mathbb{P}^k[g_{k,T} \mathbb{1}_{\{\tau_\emptyset > t\}}] = \mathbb{P}(\tau_\emptyset > t) \mathbb{P}^k[g_{k,T-t}] = e^{-rt} \mathbb{P}[N_{T-t}^{(k)}],$$

deducimos la veracidad de la primera parte del enunciado.

Para demostrar la segunda afirmación, empezamos notando que

$$\mathbb{Q}^{k,T}(\psi_1 > t) = \frac{1}{\mathbb{P}[N_T^{(k)}]} \mathbb{P}[g_{k,T} \mathbb{1}_{\{\psi_1 > t\}}],$$

mientras que la propiedad de Markov y el lema 2.1 nos permiten obtener

$$\mathbb{P}^k[g_{k,T} \mathbb{1}_{\{\psi_1 > t\}}] = \mathbb{P}^k\left[\left(\prod_{v < \xi_t^1} L_v^k\right) \mathbb{1}_{\{\psi_1 > t\}}\right] \mathbb{P}^k[g_{k,T-t}] = \mathbb{P}^k\left[\left(\prod_{v < \xi_t^1} L_v^k\right) \mathbb{1}_{\{\psi_1 > t\}}\right] \mathbb{P}[N_{T-t}^{(k)}].$$

En consecuencia

$$\mathbb{Q}^{k,T}(\psi_1 > t) = \frac{\mathbb{P}[N_{T-t}^{(k)}]}{\mathbb{P}[N_T^{(k)}]} \mathbb{P}^k\left[\left(\prod_{v < \xi_t^1} L_v^k\right) \mathbb{1}_{\{\psi_1 > t\}}\right],$$

por lo que resta tratar la esperanza que involucra la indicadora  $\mathbb{1}_{\{\psi_1 > t\}}$ . Para ello observamos que  $\psi_1 > t$  si y sólo si las  $k$  marcas especiales se encuentran asignadas al mismo individuo al tiempo  $t$ , y por ende

$$\begin{aligned} \mathbb{P}^k\left[\left(\prod_{v < \xi_t^1} L_v^k\right) \mathbb{1}_{\{\psi_1 > t\}}\right] &= \mathbb{P}^k\left[\sum_{u \in \mathcal{N}_t} \left(\prod_{v < u} L_v^k\right) \mathbb{1}_{\{\xi_t^i = u, i \in [k]\}}\right] \\ &= \mathbb{P}^k\left[\sum_{u \in \mathcal{N}_t} 1\right] = \mathbb{P}[N_t] = e^{(m-1)t}, \end{aligned}$$

donde la segunda igualdad se da aplicando un argumento análogo al de la demostración del lema 2.1, condicionando en  $\mathcal{F}_t^0$ . De lo anterior obtenemos la segunda afirmación, que junto a la primera implica inmediatamente la tercera parte del enunciado. ■

El lema 3 nos da la distribución de los tiempos de división bajo  $Q^{k,T}$ , así como nos da la tasa a la cual inmigran individuos fuera de la espina. Sin embargo, no nos da la distribución del número de inmigrantes en un evento de inmigración, el cual llamaremos *nacimiento fuera de la espina*, contenido del siguiente lema.

#### 4. Lema

Para cualquier elección  $j \geq 0$ ,  $k \in \mathbb{N}$  y  $0 \leq t < T$ ,

$$Q^{k,T}(L_\emptyset = j \mid \tau_\emptyset = t, \psi_1 > t) = \frac{jp_j}{m}.$$

Lo que el lema 4 dice es: en los nacimientos fuera de la espina, la distribución del número de descendientes es sesgada por tamaño con respecto a  $L$ . Así, el número de descendientes en un nacimiento fuera de la espina siempre es sesgado por tamaño sin importar el número de marcas especiales asignadas a una espina; esto significa que el número de marcas especiales únicamente afecta los eventos de división de espinas.

Demostración: Por la definición de  $Q^{k,T}$  y las definiciones de  $g_{k,T}$  y  $\zeta_{k,T}$ ,

$$\begin{aligned} Q^{k,T}(L_\emptyset = j \mid \tau_\emptyset = t, \psi_1 > t) &= \frac{Q^{k,T}(L_\emptyset = j, \psi_1 > t \mid \tau_\emptyset = t)}{Q^{k,T}(\psi_1 > t \mid \tau_\emptyset = t)} \\ &= \frac{\mathbb{P}^k[g_{k,T} \mathbb{1}_{\{L_\emptyset = j, \psi_1 > t\}} \mid \tau_\emptyset = t]}{\mathbb{P}^k[g_{k,T} \mathbb{1}_{\{\psi_1 > t\}} \mid \tau_\emptyset = t]}. \end{aligned}$$

En caso de que el individuo inicial tenga  $i$  descendientes, entonces en el producto que define a  $g_{k,T}$ , ver (2.1), contiene un término de la forma  $i^k$ , mientras que la probabilidad de que todas las espinas se asignen al mismo descendiente es  $1/i^{k-1}$ . Por consiguiente, debido a la propiedad de Markov bajo  $\mathbb{P}^k$ ,

$$\mathbb{P}^k[g_{k,T} \mathbb{1}_{\{L_\emptyset = i, \psi_1 > t\}} \mid \tau_\emptyset = t] = p_i i^k \frac{1}{i^{k-1}} \mathbb{P}^k[g_{k,T-t}] = ip_i \mathbb{P}[g_{k,T-t}]$$

resulta válido para toda  $i \in \mathbb{N}_0$ . Consecuentemente,

$$Q^{k,T}(L_\emptyset = j \mid \tau_\emptyset = t, \psi_1 > t) = \frac{jp_j \mathbb{P}[g_{k,T-t}]}{\sum_{i=0}^{\infty} ip_i \mathbb{P}[g_{k,T-t}]} = \frac{jp_j}{m}. \quad \blacksquare$$

Para completar la descripción del comportamiento del árbol bajo  $Q^{k,T}$ , resta entender la asignación de las marcas especiales en los eventos de división de espinas. Con este cometido recordemos que  $\mathcal{Z}_t^k$  es la partición de  $[k]$  inducida por  $\xi_t$  y  $Z_i^k = \mathcal{Z}_{\psi_i}^k$  es la partición resultante en los tiempos de división de las espinas. Debido al lema de simetría, nuevamente podemos considerar lo que sucede en el primer evento de división de espinas.

#### 5. Lema

Condicional en  $\{\psi_1 > t\}$ , la probabilidad condicional bajo  $Q^{k,T}$  de que en el intervalo  $[t, t+h)$  la espina muera y tenga  $l$  descendientes, y que las marcas se partitionen acorde a una partición  $P$  con bloques de tamaño  $a_1, \dots, a_n$  es

$$Q^{k,T}(\psi_1 < t+h, Z_1^k = P, L_{\xi_1^k} = l \mid \psi_1 > t) = p_l l^{(n)} \frac{\prod_{i=1}^n \mathbb{P}^k[N_{T-t}^{(a_i)}]}{\mathbb{P}^k[N_{T-t}^{(k)}]} (rh + o(h)).$$

Demostración: Debido al lema de simetría obtenemos la primera igualdad en

$$\begin{aligned}
& \mathbb{Q}^{k,T}(\psi_1 < t + h, Z_1^k = P, L_{\xi_1^1} = l \mid \psi_1 > t) \\
&= \mathbb{Q}^{k,T-t}(\psi_1 < h, Z_1^k = P, L_\emptyset = k) \\
&= \frac{1}{\mathbb{P}^k[N_{T-t}^{(k)}]} \mathbb{P}^k[g_{k,T-t} \mathbb{1}_{\{\psi_1 < h, Z_1^k = P, L_\emptyset = l\}}] \\
&= \frac{1}{\mathbb{P}^k[N_{T-t}^{(k)}]} \mathbb{P}^k(\psi_1 < h, Z_1^k = P, L_\emptyset = l) \mathbb{P}^k[g_{k,T-t} \mid \psi_1 < h, Z_1^k = P, L_\emptyset = l].
\end{aligned}$$

Con el desarrollo dado en mente, primero vemos que

$$\begin{aligned}
\mathbb{P}^k(\psi_1 < h, Z_1^k = P, L_\emptyset = l) &= \mathbb{P}^k(Z_1^k = P \mid \psi_1 < h, L_\emptyset = l) \mathbb{P}^k(\psi_1 < h, L_\emptyset = l) \\
&= \frac{l^{(n)}}{l^k} \mathbb{P}^k(\psi_1 < h, L_\emptyset = l),
\end{aligned}$$

con la última igualdad dándose debido a que  $l^{(n)}/l^k$  es la probabilidad de repartir  $k$  bolas de manera uniforme e independiente en  $l$  contenedores de tal forma que se induzca la partición  $P$ .

Por otra parte es claro que en el evento  $\{Z_1^k = P, L_\emptyset = l\}$ ,

$$g_{k,T-t} = l^k \prod_{p \in P} \mathbb{1}_{\{\xi_{T-t}^i \neq \xi_{T-t}^j, i,j \in p, i \neq j\}} \prod_{i \in p} \prod_{\xi_{\psi_1}^i \leq v < \xi_{T-t}^i} L_v.$$

Ahora, para cada  $p \in P$ , el lema 2.1 implica que en el conjunto  $\{Z_1^k = P, L_\emptyset = l\}$ ,

$$\mathbb{P}^k \left[ \mathbb{1}_{\{\xi_{T-t}^i \neq \xi_{T-t}^j, i,j \in p, i \neq j\}} \prod_{i \in p} \prod_{\xi_{\psi_1}^i \leq v < \xi_{T-t}^i} L_v \mid \mathcal{F}_{\psi_1}^k \right] = \mathbb{P}^k[N_{T-t-u}^{(\#p)}] \Big|_{u=\psi_1},$$

y en el evento  $\{\psi_1 < h\}$  se cumple

$$\mathbb{P}^k[N_{T-t-u}^{(\#p)}] \Big|_{u=\psi_1} = \mathbb{P}^k[N_{T-t}^{(\#p)}] + o(1),$$

de modo que

$$\mathbb{P}^k[g_{k,T-t} \mid \psi_1 < h, Z_1^k = P, L_\emptyset = l] = l^k \left( \prod_{i=1}^n \mathbb{P}^k[N_{T-t}^{(a_i)}] + o(1) \right).$$

Al juntar los resultados obtenidos hasta el momento,

$$\begin{aligned}
& \mathbb{Q}^{k,T}(\psi_1 < t + h, Z_1^k = P, L_{\xi_1^1} = l \mid \psi_1 > t) \\
&= \frac{1}{\mathbb{P}^k[N_{T-t}^{(k)}]} \times p_l \frac{l^{(n)}}{l^k} (rh + o(h)) \times l^k \left( \prod_{i=1}^n \mathbb{P}^k[N_{T-t}^{(a_i)}] + o(1) \right) \\
&= p_l l^{(n)} \frac{\prod_{i=1}^n \mathbb{P}^k[N_{T-t}^{(a_i)}]}{\mathbb{P}^k[N_{T-t}^{(k)}]} (rh + o(h)),
\end{aligned}$$

siendo esto lo que se buscaba demostrar. ■

Para cerrar la sección haremos una última observación. Dada una colección de enteros positivos  $a_1, \dots, a_n$  cuya suma sea  $k$ , si denotamos por

$$n_j = \#\{i : a_i = j\},$$

entonces  $\sum_{i=1}^k n_i = n$  y  $\sum_{i=1}^k i n_i = k$ . Luego, el número de particiones de  $[k]$  en bloques de tamaños  $a_1, \dots, a_n$ , es

$$\frac{k!}{\prod_{i=1}^n a_i!} \frac{1}{\prod_{j=1}^k n_j!}.$$

Al combinar este resultado con los lemas 3 y 5 deducimos el siguiente resultado, con el cual concluimos el apartado.

### 6. Corolario

Se cumple que

$$\begin{aligned} Q^{k,T}(\psi_1 \in [t, t + dt), L_{\xi_t^1} = l, \#\{p \in Z_1^k : \#p = j\} = n_j \text{ para } j \in [k]) \\ = \frac{l^{(n)} p_l}{\mathbb{P}[L^{(n)}]} \frac{k!}{\prod_{i=1}^n a_i! \prod_{j=1}^{k-1} n_j!} \mathbb{P}[L^{(n)}] r e^{(m-1)rt} \frac{\prod_{i=1}^n \mathbb{P}^k[N_{T-t}^{(a_i)}]}{\mathbb{P}^k[N_T^{(k)}]} dt. \end{aligned}$$

## 7.4. Comportamiento de $Q^{k,T}$ para tiempos grandes

En la presente sección estudiaremos el comportamiento del árbol dado por las espinas cuando consideramos el límite conforme  $T \rightarrow \infty$ , para tener un mejor entendimiento de los resultados presentados en la sección 7.2. Para ser sucintos únicamente nos enfocaremos en el caso crítico donde  $m = 1$ .

Tomemos  $m = 1$ ,  $n = 2$  y  $t = sT$  en el corolario 3.6. En el caso  $a_1 \neq a_2$  se tiene, considerando  $E = \{\text{las espinas se dividen en dos grupos de tamaños } a_1 \text{ y } a_2\}$ ,

$$\begin{aligned} Q^{k,T}(\psi_1 \in [sT, sT + T ds), L_{\xi_{sT}^1} = l, E) \\ = \frac{l(l-1)p_l}{\mathbb{P}^k[L(L-1)]} \frac{k!}{a_1! a_2!} \mathbb{P}^k[L(L-1)] r \frac{\mathbb{P}^k[N_{T(1-s)}^{(a_1)}] \mathbb{P}^k[N_{T(1-s)}^{(a_2)}]}{\mathbb{P}^k[N_T^{(k)}]} T ds. \end{aligned}$$

Así, al tomar el límite conforme  $T \rightarrow \infty$  podemos usar el teorema VI.2.1 para deducir que  $T\mathbb{P}(N_{uT} > 0) \rightarrow 2/(\sigma^2 r u)$  conforme  $T \rightarrow \infty$  y, condicional en  $\{N_{uT} > 0\}$ ,  $N_{uT}/T$  converge a una variable aleatoria con distribución exponencial de parámetro  $2/(\sigma^2 r u)$ . Consecuentemente,

$$\begin{aligned} \lim_{T \rightarrow \infty} Q^{k,T}(\psi_1 \in [sT, sT + T ds), L_{\xi_{sT}^1} = l, E) \\ = \frac{l(l-1)p_l}{\mathbb{P}[L(L-1)]} \frac{k!}{a_1! a_2!} \mathbb{P}[L(L-1)] r \frac{\frac{2}{\sigma^2 r(1-s)} \left(\frac{\sigma^2 r(1-s)}{2}\right)^{a_1} a_1! \frac{2}{\sigma^2 r(1-s)} \left(\frac{\sigma^2 r(1-s)}{2}\right)^{a_2} a_2!}{\frac{2}{\sigma^2 r} \times \left(\frac{\sigma^2 r}{2}\right)^k k!} ds \\ = l(l-1)p_l r \frac{2}{\sigma^2} (1-s)^{k-2} ds. \end{aligned}$$

En caso de que  $a_1 = a_2$ , es necesario incluir un factor de  $1/2$  debido a que los dos bloques se pueden reordenar.

Ahora, debido a que hay  $k-1$  maneras en las cuales se pueden dividir las  $k$  espinas en dos grupos de tamaño mayor que cero, y de que lo anterior nos dice que todos

los resultados son igual de probables,

$$\begin{aligned} \lim_{T \rightarrow \infty} \mathbb{Q}^{k,T}(\psi_1 \in [sT, sT + T ds), L_{\xi_{sT}}^1 = l, \text{ las espinas se dividen en dos grupos}) \\ = \frac{l(l-1)}{\sigma^2} p_l(k-1)(1-s)^{k-2} ds. \end{aligned}$$

Se puede observar que al sumar la expresión anterior sobre  $l$ , y luego integrar sobre  $s \in [0, 1]$  obtenemos 1. Esto quiere decir que en el límite, conforme  $T \rightarrow \infty$ , en el primer evento de división  $\psi_1$ , las  $k$  espinas siempre se separan en exactamente dos grupos. Más aún, el número de espinas en cada uno de los grupos es uniforme en  $\{1, \dots, k-1\}$ , y el número total de descendientes en este tiempo se encuentra doblemente sesgada por tamaño. En adición, el primer tiempo de división, reescalado por  $T^{-1}$ , converge en distribución al mínimo de  $k$  variables aleatorias independientes con distribución uniforme en  $[0, 1]$ .

En vista del lema de simetría, lema 3.2, nos dice que podemos extender nuestro entendimiento del primer tiempo de división a todos los tiempos de división de espinas. Así, cuando una espina se divide ella, en el límite conforme  $T \rightarrow \infty$ , se divide en dos grupos de manera uniforme. Tal propiedad es compartida con el árbol que surge del árbol asociado a un coalescente de Kingman. Adicionalmente, los  $k-1$  tiempos de división, reescalados por  $T^{-1}$  son independientes y se distribuyen de manera uniforme en  $[0, 1]$ .

Lo dicho en este apartado solamente es cierto para el caso crítico. De estar interesados en el caso casi-crítico, cuando  $m = 1 + \mu/T + o(1/T)$ , entonces la densidad uniforme para los tiempos de división es reemplazada  $(e^{r\mu} - 1)^{-1} r\mu e^{r\mu s} ds$ . Esto quiere decir que el caso casi-crítico es un cambio de tiempo determinista del caso crítico.

## 7.5. Demostración de la proposición 2.2

## 7.6. Fórmula de Campbell

### 1. Proposición

Para cada  $z \geq 0$ ,

$$\mathbb{Q}^{k,T}[e^{-z\tilde{N}_T} \mid \mathcal{G}_T^k] = \prod_{i=0}^{k-1} \left( e^{-r(m-1)(T-\psi_i)} \frac{u(F(e^{-z}, T - \psi_i))}{u(e^{-z})} \right)$$

$\mathbb{Q}^{k,T}$ -casi seguramente.

# Procesos de nacimiento y muerte

---

8

## Límite de escala en el caso casi-crítico

---

9

## Discusión de los resultados

---

# 10

## **Parte III.**

# **Apéndices**

# Cálculo estocástico



El objetivo de este capítulo es recopilar algunos resultados del tiempo local, el cual se introdujo en la sección 2.2 para procesos de Markov, desde la perspectiva del cálculo estocástico. Esto en particular se aplicará al movimiento browniano, lo cual es de utilidad en el cuerpo principal de las notas. Lo expuesto en el capítulo se basa en el capítulo VI de Revuz y Yor (1999).

## A.1. Tiempo local para semimartingalas

La fórmula de Itô nos dice cómo es que operan las funciones de clase  $C^2$  en semimartingalas continuas. Esta idea se puede extender a las funciones convexas, lo que permite introducir la noción de tiempo local desde el punto de vista de cálculo estocástico. A lo largo de la sección consideraremos que  $f$  es una función convexa.

### 1. Teorema

Si  $X$  es una semimartingala continua, existe un proceso creciente y continuo  $A^f$  tal que

$$f(X_t) = f(X_0) + \int_0^t f'_-(X_s) dX_s + \frac{1}{2}A_t^f,$$

donde  $f'_-$  es la derivada por la izquierda de  $f$ .

Demostración: Si  $f \in C^2(\mathbb{R})$  el resultado se deduce de la fórmula de Itô poniendo  $A_t^f = \int_0^t f''(X_s) d\langle X \rangle_s$ .

Consideremos una función  $j \in C_c^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$  con soporte en  $(-\infty, 0]$  tal que  $\int_{-\infty}^0 j(x) dx = 1$  y definamos  $f_n(x) = n \int_{-\infty}^0 f(x+y)j(ny) dy$ , función que está bien definida pues  $f$  es localmente acotada. Más aún,  $f_n \rightarrow f$  puntualmente conforme  $n \rightarrow \infty$ , mientras que  $f'_n \not\rightarrow f'_-$ . Por la fórmula de Itô, para cada  $n$  tenemos

$$f_n(X_t) = f_n(X_0) + \int_0^t f'_n(X_s) dX_s + \frac{1}{2}A_t^{f_n},$$

y  $f_n(X_t)$  converge a  $f(X_t)$ . Ahora, podemos suponer que  $X$  es acotado, implicando que  $f'_-(X_s)$  es también acotado. Así, por el teorema de convergencia dominada para integrales estocásticas,

$$\int_0^t f'_n(X_s) dX_s \rightarrow \int_0^t f'_-(X_s) dX_s$$

en probabilidad de manera uniforme en compactos. Por ende  $A^{f_n}$  converge a un proceso  $A^f$  que es creciente por ser límite de procesos crecientes y en adición satisface

$$f(X_t) = f(X_0) + \int_0^t f'_-(X_s) dX_s + \frac{1}{2}A_t^f.$$

Basta escoger a  $A_t^f$  de tal forma que sea casi seguramente continuo para concluir. ■

En general es complicado calcular  $A^f$  de tal suerte que la expresión sea explícita y útil para hacer claro cómo es que depende de  $f$ . El contenido del siguiente teorema consiste en estos cálculos para los casos especiales de  $|\cdot|$ ,  $x_+$  y  $x_-$ .

## 2. Teorema (Fórmula de Tanaka)

Para cada número real  $a$  existe un proceso continuo y creciente  $L^a$  para el cual se cumplen

$$\begin{aligned} |X_t - a| &= |X_0 - a| + \int_0^t \operatorname{sgn}(X_s - a) dX_s + L_t^a \\ (X_t - a)_+ &= (X_0 - a)_+ + \int_0^t \mathbb{1}_{\{X_s > a\}} dX_s + \frac{1}{2} L_t^a \\ (X_t - a)_- &= (X_0 - a)_- - \int_0^t \mathbb{1}_{\{X_s \leq a\}} dX_s + \frac{1}{2} L_t^a. \end{aligned}$$

En particular deducimos que  $|X - a|$ ,  $(X - a)_+$  y  $(X - a)_-$  son semimartingalas.

Demostración: Puesto que la derivada por la izquierda de  $f(x) = (x - a)_+$  es  $\mathbb{1}_{(a, \infty)}$ , el teorema 1 implica que existe un proceso  $A^+$  tal que

$$(X_t - a)_+ = (X_0 - a)_+ + \int_0^t \mathbb{1}_{\{X_s > a\}} dX_s + \frac{1}{2} A_t^+.$$

De manera análoga

$$(X_t - a)_- = (X_0 - a)_- - \int_0^t \mathbb{1}_{\{X_s \leq a\}} dX_s + \frac{1}{2} A_t^-.$$

Consecuentemente

$$X_t = X_0 + \int_0^t dX_s + \frac{1}{2} (A_t^+ - A_t^-),$$

de donde se sigue que  $A_t^+ = A_t^-$  casi seguramente. Concluimos al poner  $L_t^a = A_t^+$ . ■

## 3. Proposición

El soporte de la medida  $dL_t^a$  se encuentra contenido en  $\{t \in \mathbb{R} : X_t = a\}$  con probabilidad uno.

Demostración: Usando la fórmula de Itô y la fórmula de Tanaka obtenemos

$$\begin{aligned} (X_t - a)^2 &= (X_0 - a)^2 + 2 \int_0^t |X_s - a| d|X_s - a| + \langle |X - a| \rangle_t \\ &= (X_0 - a)^2 + 2 \int_0^t |X_s - a| \operatorname{sgn}(X_s - a) dX_s + 2 \int_0^t |X_s - a| dL_s^a + \langle X \rangle_t. \end{aligned}$$

Por otra parte, la fórmula de Itô nos da directamente

$$(X_t - a)^2 = (X_0 - a)^2 + 2 \int_0^t (X_s - a) dX_s + \langle X \rangle_t.$$

Comparando ambas expresiones deducimos que  $\int_0^t |X_s - a| dL_s^a = 0$  casi seguramente, lo cual nos da el resultado. ■

Este resultado y la proposición II.2.9 nos dicen que el proceso  $L^a$  coincide, salvo constantes multiplicativas, con el tiempo local definido para procesos de Markov en general. Por ende llamaremos a  $L^a$  el *tiempo local de X en a*.

**4. Lema**

Existe un proceso  $\tilde{L}$ , medible con respecto a  $\mathcal{B}(\mathbb{R}) \otimes \mathcal{P}$ , tal que  $\tilde{L}(a, \cdot, \cdot)$  es indistinguible de  $L^a$  para cada  $a \in \mathbb{R}$ .

Demostración: El resultado se sigue al aplicar el teorema de Fubini para integrales estocásticas al proceso  $H(a, s, \cdot) = \mathbb{1}_{\{X_s > a\}}$ . ■

**5. Teorema** (Fórmula de Itô–Tanaka)

Si  $f$  es la diferencia de dos funciones convexas y si  $X$  es una semimartingala continua,

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_{\mathbb{R}} L_t^a df''(a).$$

En particular,  $f$  es una semimartingala.

Demostración: Es suficiente demostrar que la fórmula es cierta cuando  $f$  es convexa. Debido a que en cada compacto  $f$  coincide con una función convexa  $g$  con la propiedad de que  $g''$  tiene soporte compacto, podemos suponer que  $f''$  tiene soporte compacto, deteniendo  $X$  al salir de este. Bajo este supuesto existen dos constante  $\alpha$  y  $\beta$  tales que

$$f(x) = \alpha x + \beta + \frac{1}{2} \int_{\mathbb{R}} |x - a| df''(a).$$

De esta igualdad y la fórmula de Tanaka,

$$\begin{aligned} f(X_t) &= \alpha X_t + \beta + \frac{1}{2} \int_{\mathbb{R}} |X_t - a| df''(a) \\ &= \alpha(X_t - X_0) + \frac{1}{2} \int_{\mathbb{R}} (|X_t - a| - |X_0 - a|) df''(a) \\ &= \alpha(X_t - X_0) + \frac{1}{2} \int_{\mathbb{R}} \left( \int_0^t \operatorname{sgn}(X_s - a) dX_s + L_t^a \right) df''(a). \end{aligned}$$

Ahora, por el lema 4 y propiedades de las funciones convexas,

$$\frac{1}{2} \int_{\mathbb{R}} \int_0^t \operatorname{sgn}(X_s - a) dX_s df''(a) = \int_0^t f'_-(X_s) dX_s - \alpha(X_t - X_0),$$

igualdad que nos permite finalizar la demostración. ■

**6. Corolario** (Fórmula de tiempos de ocupación)

Existe un conjunto  $\mathbb{P}$ -nulo fuera del cual

$$\int_0^t \Phi(X_s) d\langle X \rangle_s = \int_{-\infty}^{\infty} \Phi(a) L_t^a da$$

para cualquier  $t$  y toda función  $\Phi : \mathbb{R} \rightarrow \mathbb{R}_+$  medible.

Demostración: Si  $\Phi = f''$  con  $f \in C^2(\mathbb{R})$  la igualdad, para toda  $t$ , es una consecuencia de comparar la fórmula de Itô con la de Itô–Tanaka, fuera de un conjunto  $\Gamma_\Phi$   $\mathbb{P}$ -nulo. Para extender el resultado al caso general, basta considerar una colección contable  $\{\Phi_n\} \subset C^2(\mathbb{R})$  densa en  $C_0(\mathbb{R})$  para la topología de convergencia uniforme. En efecto, en tal caso vemos que la igualdad se da de manera simultánea para

cualesquiera  $t$  y  $\Phi \in C_0(\mathbb{R})$  fuera del conjunto  $\mathbb{P}$ -nulo  $\Gamma = \bigcup_n \Gamma_{\Phi_n}$ . Resta aplicar el teorema de clases monótonas para finalizar la prueba. ■

### 7. Teorema

Para cada semimartingala continua  $X$  existe una modificación de  $\{L_t^a : a \in \mathbb{R}, t \in \mathbb{R}_+\}$  que satisface que el mapeo  $(a, t) \mapsto L_t^a$  es casi seguramente continuo en  $t$  y càdlàg en  $a$ . En adición, si  $X = M + V$ , entonces

$$L_t^a - L_t^{a-} = 2 \int_0^t \mathbb{1}_{\{X_s=a\}} dV_s = 2 \int_0^t \mathbb{1}_{\{X_s=a\}} dX_s.$$

Consecuentemente, si  $X$  es una martingala local, existe una modificación bicontinua de la familia de tiempos locales  $\{L^a : a \in \mathbb{R}\}$ .

Demostración: De la fórmula de Tanaka tenemos que

$$(1) \quad L_t^a = 2[(X_t - a)_+ - (X_0 - a)_+ - \int_0^t \mathbb{1}_{\{X_s > a\}} dM_s - \int_0^t \mathbb{1}_{\{X_s > a\}} dV_s].$$

Usaremos el criterio de continuidad de Kolmogorov para demostrar que la integral estocástica

$$\tilde{M}_t^a = \int_0^t \mathbb{1}_{\{X_s > a\}} dM_s$$

tiene una modificación bicontinua.

En virtud de las desigualdades de Burkholder–Davis–Gundy y el corolario 6, para cualquier  $k \geq 1$ ,

$$\begin{aligned} \mathbb{E}\left[\sup_t |\tilde{M}_t^a - \tilde{M}_t^b|^{2k}\right] &\leq C_k \mathbb{E}\left[\left(\int_0^\infty \mathbb{1}_{\{a < X_s \leq b\}} d\langle M \rangle_s\right)^k\right] \\ &= C_k \mathbb{E}\left[\left(\int_a^b L_\infty^x dx\right)^k\right] \\ &= C_k (b-a)^k \mathbb{E}\left[\left(\frac{1}{b-a} \int_a^b L_\infty^x dx\right)^k\right] \\ &\leq C_k (b-a)^k \mathbb{E}\left[\frac{1}{b-a} \int_a^b (L_\infty^x)^k dx\right]. \end{aligned}$$

Al aplicar Fubini obtenemos que

$$C_k (b-a)^k \mathbb{E}\left[\frac{1}{b-a} \int_a^b (L_\infty^x)^k dx\right] \leq C_k (b-a)^k \sup_x \mathbb{E}[(L_\infty^x)^k].$$

De (1) y que  $(\cdot)_+$  es Lipschitz de constante uno deducimos la existencia de una constante universal  $d_k$  para la cual se cumple

$$\mathbb{E}[(L_\infty^x)^k] \leq d_k \mathbb{E}\left[\sup_t |X_t - X_0|^k + \left(\int_0^\infty |dV_s|\right)^k + \langle M \rangle_\infty^{k/2}\right],$$

cota que no depende de  $x$ . En el caso de que esta cota sea finita para algún valor de  $k > 1$  obtenemos la afirmación. En caso contrario, paramos  $X$  en los tiempos

$$T_n = \inf\left\{t : \sup_{s \leq t} |X_s - X_0|^k + \left(\int_0^t |dV_s|\right)^k + \langle M \rangle_t^{k/2} \geq n\right\}.$$

Puesto que las martingalas  $(\tilde{M}^a)^{T_n}$  tienen modificaciones bicontinuas, también  $\tilde{M}^a$  la tendrá.

Para concluir resta probar que

$$\tilde{V}_t^a = \int_0^t \mathbb{1}_{\{X_s > a\}} dV_s$$

es conjuntamente continua en  $t$  y càdlàg en  $a$ . Por el teorema de Lebesgue,

$$\tilde{V}_t^{a-} = \lim_{b \rightarrow a^-} \int_0^t \mathbb{1}_{\{X_s > b\}} dV_s = \int_0^t \mathbb{1}_{\{X_s \geq a\}} dV_s,$$

de donde se sigue que

$$L_t^a - L_t^{a-} = 2(\tilde{V}_t^{a-} - \tilde{V}_t^a) = 2 \int_0^t \mathbb{1}_{\{X_s = a\}} dV_s.$$

De manera similar vemos que

$$\tilde{V}_t^{a+} = \lim_{b \rightarrow a^+} \int_0^t \mathbb{1}_{\{X_s > b\}} dV_s = \int_0^t \mathbb{1}_{\{X_s > a\}} dV_s,$$

de modo que  $L_t^a = L_t^{a+}$ .

Por último, la fórmula de ocupación implica

$$\int_0^t \mathbb{1}_{\{X_s = a\}} d\langle M \rangle_s = \int_0^t \mathbb{1}_{\{X_s = a\}} d\langle X \rangle_s = 0,$$

y por tanto  $\int_0^t \mathbb{1}_{\{X_s = a\}} dM_s = 0$ , finalizando la demostración. ■

Continuaremos usando la versión del tiempo local dada por el teorema 7 de ahora en adelante.

## 8. Corolario

Si  $X$  es una semimartingala continua entonces, con probabilidad uno,

$$L_t^a(X) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^t \mathbb{1}_{[a, a+\varepsilon)}(X_s) d\langle X \rangle_s$$

para toda elección de  $a$  y  $t$ . Por otra parte, si  $M$  es una martingala local continua,

$$L_t^a(M) = \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{(a-\varepsilon, a+\varepsilon)}(M_s) d\langle M \rangle_s.$$

El resultado se sigue satisfaciendo si se cambia  $t$  por cualquier tiempo aleatorio  $S$ .

Demostración: Esto es una consecuencia de la fórmula de ocupación y la continuidad por la derecha en  $a$  de  $L_t^a(X)$ . ■

De manera particular, para el movimiento browniano se cumple que

$$L_t^0 = \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{\{|B_s| < \varepsilon\}} ds,$$

lo cual retomaremos en el apéndice A.2.

El corolario 8 es un resultado de “aproximación” para el tiempo local, misma que puede usarse para definirlo en otros contextos, ver por ejemplo el capítulo v

de Bertoin (1996). A continuación veremos otro resultado de aproximación que es válido para una gran clase de semimartingalas. Dada  $X$  una semimartingala continua definimos, considerando  $\varepsilon > 0$  fijo, la sucesión de tiempos de paro  $\sigma_0^\varepsilon < \tau_0^\varepsilon < \sigma_1^\varepsilon < \dots$  definida por  $\sigma_0^\varepsilon = 0$ , mientras que para  $n \in \mathbb{N}_0$  ponemos

$$\tau_n^\varepsilon = \inf\{t > \sigma_n^\varepsilon : X_t = \varepsilon\} \quad \text{y} \quad \sigma_{n+1}^\varepsilon = \inf\{t > \tau_n^\varepsilon : X_t = 0\}.$$

Ponemos  $N_t^\varepsilon = \max\{n : \sigma_n^\varepsilon < t\}$ , que en la literatura de cálculo estocástico se conoce como el número de *downcrossings* de  $X$  del nivel  $\varepsilon$  al nivel 0 antes del tiempo  $t$ . Retomando las ideas del capítulo 2,  $N_t^\varepsilon$  cuenta el número de excursiones fuera de cero que alcanzaron un nivel  $\varepsilon$  completadas antes del tiempo  $t$ .

El siguiente resultado nos dirá que podemos aproximar al tiempo local mediante una renormalización adecuada de  $N_t^\varepsilon$ . Para simplificar la notación escribiremos  $\tau_n$ ,  $\sigma_n$  y  $L_t$  en lugar de  $\tau_n^\varepsilon$ ,  $\sigma_n^\varepsilon$  y  $L_t^0$  de forma respectiva.

### 9. Teorema

Sea  $X = M + V$  una semimartingala tal que

$$\mathbb{E}\left[\langle M \rangle_\infty^{p/2} + \left(\int_0^\infty |dV|_s\right)^p\right] < \infty$$

para alguna  $p \geq 1$ . Entonces

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}\left[\sup_t \left|\varepsilon N_t^\varepsilon - \frac{1}{2}L_t\right|^p\right] = 0.$$

Demostración: De la fórmula de Tanaka obtenemos

$$X_{\tau_n \wedge t}^+ - X_{\sigma_n \wedge t}^+ = \int_{(\sigma_n \wedge t, \tau_n \wedge t]} \mathbb{1}_{\{X_s > 0\}} dX_s + \frac{1}{2}(L_{\tau_n \wedge t} - L_{\sigma_n \wedge t}).$$

Puesto que  $X$  es distinta de cero en  $[\tau_n, \sigma_n)$  vemos que  $L_{\tau_n \wedge t} - L_{\sigma_n \wedge t} = L_{\sigma_{n+1} \wedge t} - L_{\sigma_n \wedge t}$  y como consecuencia

$$\sum_n (X_{\tau_n \wedge t}^+ - X_{\sigma_n \wedge t}^+) = \int_0^t \Psi_s^\varepsilon dX_s + \frac{1}{2}L_t,$$

donde  $\Psi_s^\varepsilon$  es el proceso predecible  $\sum_n \mathbb{1}_{(\sigma_n, \tau_n]}(s) \mathbb{1}_{(0, \varepsilon]}(X_s)$ . Como en  $\{\tau_n \leq t\}$  tenemos que  $X_{\tau_n \wedge t}^+ - X_{\sigma_n \wedge t}^+ = \varepsilon$ , si  $n(t) = \inf\{n : \tau_n > t\}$ ,

$$\sum_n (X_{\tau_n \wedge t}^+ - X_{\sigma_n \wedge t}^+) = \varepsilon N_t^\varepsilon + u(\varepsilon),$$

donde  $0 \leq u(\varepsilon) = X_t^+ - X_{\sigma_{n(t)} \wedge t} \leq \varepsilon$ . Así, será suficiente demostrar que

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}\left[\sup_t \left|\int_0^t \Psi_s^\varepsilon dX_s\right|^p\right] = 0$$

para concluir con la demostración. En esta dirección, observamos que las desigualdades de Burkholder–Davis–Gundy implican que la esperanza previa se encuentra acotada por

$$C_p \mathbb{E}\left[\left(\int_0^\infty \Psi_s^\varepsilon d\langle X \rangle_s\right)^{p/2} + \left(\int_0^\infty \Psi_s^\varepsilon |dV|_s\right)^p\right].$$

Debido a que  $\Psi^\varepsilon$  converge de forma acotada a cero, finalizamos mediante una aplicación del teorema de convergencia dominada. ■

## A.2. Tiempo local para el movimiento browniano

A lo largo de este apartado estudiaremos algunas de las propiedades de los tiempos locales de un movimiento browniano estándar. De los resultados vistos en la apéndice A.1 sabemos que hay una familia bicontinua  $\{L_t^a : a \in \mathbb{R}, t \in \mathbb{R}_+\}$  de tiempos locales. Debido a que nuestro interés será en particular sobre el tiempo local en 0, escribiremos  $L_t$  en lugar de  $L_t^0$ .

Siguiendo parte de la notación del capítulo 2, para  $t \geq 0$ ,

$$R_t = \inf\{s > t : X_s = 0\}$$

será el primer tiempo de retorno al cero después del tiempo  $t$ . Adicionalmente,  $\mathcal{Z}$  es el conjunto de ceros de  $X$ . Consecuentemente,  $L_t = L_{R_t}$ .

Lo primero que haremos será encontrar la ley del proceso  $L$ . Para ello usaremos la fórmula de Tanaka, la cual nos da la igualdad

$$|B_t| = \int_0^t \text{sgn}(B_s) dB_s + L_t.$$

Es sabido que  $|B|$  es un proceso de Markov y por lo anterior es una semimartingala cuyo tiempo local en 0 es igual a  $2L$ . Un proceso que tenga la misma ley que  $|B|$  será llamado *movimiento browniano reflejado*. Para analizar  $L$  es necesario el siguiente lema.

### 1. Lema (Skorokhod)

Sea  $y : \mathbb{R}_+ \rightarrow \mathbb{R}$  una función continua tal que  $y(0) = 0$ . Existe un único par de funciones  $(z, a)$  en  $\mathbb{R}_+$  tales que

- (i)  $z = y + a$ ,
- (ii)  $z$  es positiva y
- (iii)  $a$  es creciente, continua, se desvanece en cero y la medida  $da$  tiene soporte en  $z^{-1}(0)$ .

La función  $a$  está explícitamente dada por

$$a(t) = \sup_{s \leq t} y_-(s).$$

Demostración: Es claro que las funciones

$$a(t) = \sup_{s \leq t} y_-(s) \quad \text{y} \quad z = y + a$$

satisfacen las propiedades requeridas en el enunciado. Luego, basta probar la unicidad de la pareja  $(z, a)$ .

Supongamos pues que  $(\tilde{z}, \tilde{a})$  es una pareja de funciones que satisfacen las tres propiedades. Entonces  $z - \tilde{z} = a - \tilde{a}$  es una función de variación acotada, lo cual nos permite usar integración por partes para obtener

$$0 \leq (z(t) - \tilde{z}(t))^2 = 2 \int_0^t (z(s) - \tilde{z}(s)) d(a - \tilde{a})(s).$$

Ahora bien, por la propiedad (iii) el último término es igual a

$$-2 \int_0^t \tilde{z}(s) da(s) - 2 \int_0^t z(s) d\tilde{a}(s).$$

Las propiedades (ii) y (iii) nos dice que este término es menor o igual que 0, y por consiguiente  $0 \leq (z(t) - \tilde{z}(t))^2 \leq 0$ , o bien  $z = \tilde{z}$  y  $a = \tilde{a}$ . ■

## 2. Corolario

El proceso  $\beta_t = \int_0^t \operatorname{sgn}(B_s) dB_s$  es un movimiento browniano estándar y  $\mathcal{F}_t^\beta = \mathcal{F}_t^{|B|}$ . En adición  $L_t = \sup_{s \leq t} -\beta_s$ .

Demostración: El hecho de que  $\beta$  es un movimiento browniano estándar es una consecuencia inmediata del teorema de caracterización de Lévy. Mientras tanto, la fórmula de Tanaka nos dice que  $|B| = \beta + L$ . Tal igualdad y el lema 1 implican la igualdad  $L_t = \sup_{s \leq t} -\beta_s$ . De lo anterior también es claro que  $\mathcal{F}_t^{|B|} \subset \mathcal{F}_t^\beta$ . Por otra parte, el corolario 1.8 implica  $\mathcal{F}_t^L \subset \mathcal{F}_t^{|B|}$  y por ende  $\mathcal{F}_t^\beta \subset \mathcal{F}_t^{|B|}$ . ■

## 3. Teorema (Identidad de Lévy)

Los procesos  $(S - B, S)$  y  $(|B|, L)$  tienen la misma ley.

Demostración: Por una parte,  $|B| = \beta + L$  por la fórmula de Tanaka; por otra parte,  $S - B = -B + S$  de manera trivial. Así, por el lema 1, tomando  $y = \beta$  obtenemos la pareja  $(|B|, L)$  y tomando  $y = -B$  obtenemos  $(S - B, S)$ . El resultado se sigue entonces de que  $-B$  y  $\beta$  tienen la misma ley. ■

# Referencias

---

## Principales

- Harris, S. C., Johnston, S. G. G., & Roberts, M. I. (2020). The Coalescent Structure of Continuous-Time Galton–Watson Trees. *The Annals of Applied Probability*, 30(3). <https://doi.org/10.1214/19-AAP1532>
- Le Gall, J.-F. (2005). Random Trees and Applications. *Probability Surveys*, 2. <https://doi.org/10.1214/154957805100000140>
- Le Gall, J.-F., & Le Jan, Y. (1998). Branching Processes in Lévy Processes: The Exploration Process. *The Annals of Probability*, 26(1). <https://doi.org/10.1214/aop/1022855417>
- Pardo Millán, J. C. (s.f.-a). *Local Times and Excursion Theory* [Lecture Notes].
- Pardo Millán, J. C. (s.f.-b). *Tiempos locales y excursiones brownianas* [Lecture Notes].

## Suplementarias

- Aigner, M., & Ziegler, G. M. (2018). *Proofs from THE BOOK*. Springer. <https://doi.org/10.1007/978-3-662-57265-8>
- Aldous, D. (1991a). The Continuum Random Tree II: An Overview. En M. T. Barlow & N. H. Bingham (Eds.), *Stochastic Analysis: Proceedings of the Durham Symposium on Stochastic Analysis, 1990* (pp. 23-70). Cambridge University Press. <https://doi.org/10.1017/CBO9780511662980.003>
- Aldous, D. (1991b). The Continuum Random Tree. I. *The Annals of Probability*, 19(1), 1-28. <https://www.jstor.org/stable/2244250>
- Aldous, D. (1993). The Continuum Random Tree III. *The Annals of Probability*, 21(1), 248-289. Consultado el 30 de agosto de 2023, desde <https://www.jstor.org/stable/2244761>
- Athreya, K. B., & Ney, P. E. (1972). *Branching Processes*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-65371-1>
- Belkin, B. (1972). An invariance principle for conditioned recurrent random walk attracted to a stable law. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 21(1), 45-64. <https://doi.org/10.1007/BF00535106>
- Bertoin, J. (1996). *Lévy Processes*. Cambridge University Press.
- Billingsley, P. (1995). *Probability and Measure* (3.<sup>a</sup> ed.). Wiley.
- Billingsley, P. (1999). *Convergence of Probability Measures* (2.<sup>a</sup> ed.). Wiley.
- Blumenthal, R. M., & Gettoor, R. K. (1968). *Markov Processes and Potential Theory* (1.<sup>a</sup> ed.) [2nd print]. Academic Press.
- Burago, D., Burago, I. D., & Ivanov, S. (2001). *A Course in Metric Geometry*. American Mathematical Society.
- Chiswell, I. (2001). *Introduction to  $[\Lambda]$ -Trees*. World Scientific.

- Davis, B., & McDonald, D. (1995). An elementary proof of the local central limit theorem. *Journal of Theoretical Probability*, 8(3), 693-701. <https://doi.org/10.1007/BF02218051>
- Ethier, S. N., & Kurtz, T. G. (2005). *Markov Processes: Characterization and Convergence*. Wiley Interscience.
- Evans, S. N., Pitman, J., & Winter, A. (2006). Rayleigh processes, real trees, and root growth with re-grafting. *Probability Theory and Related Fields*, 134(1), 81-126. <https://doi.org/10.1007/s00440-004-0411-6>
- Evans, S. N., & Winter, A. (2006). Subtree Prune and Regraft: A Reversible Real Tree-Valued Markov Process. *The Annals of Probability*, 34(3). <https://doi.org/10.1214/009117906000000034>
- Ito, K., & McKean, H. P. (1996). *Diffusion Processes and Their Sample Paths*. Springer.
- Itô, K. (1972). Poisson Point Processes Attached to Markov Processes. En *Poisson Point Processes Attached to Markov Processes* (pp. 225-240). University of California Press. <https://doi.org/10.1525/9780520375918-015>
- Kaigh, W. D. (1976). An Invariance Principle for Random Walk Conditioned by a Late Return to Zero. *The Annals of Probability*, 4(1), 115-121. Consultado el 4 de abril de 2023, desde <https://www.jstor.org/stable/2959376>
- Kechris, A. S. (1995). *Classical Descriptive Set Theory*. Springer-Verlag.
- Kesten, H. (1963). Ratio theorems for random walks II. *Journal d'Analyse Mathématique*, 11(1), 323-379. <https://doi.org/10.1007/BF02789990>
- Le Gall, J.-F. (2000). Random Trees And Spatial Branching Processes. <https://www.semanticscholar.org/paper/Random-Trees-And-Spatial-Branching-Processes-Gall/1689e9c2403ff6496701c9ee94df6ff26c8a0487>
- Le Gall, J.-F. (2010). Itô's excursion theory and random trees. *Stochastic Processes and their Applications*, 120(5), 721-749. <https://doi.org/10.1016/j.spa.2010.01.015>
- Le Gall, J.-F. (2016). *Brownian Motion, Martingales, and Stochastic Calculus* (Vol. 274). Springer International Publishing. <https://doi.org/10.1007/978-3-319-31089-3>
- Lyons, R., Pemantle, R., & Peres, Y. (1995). Conceptual Proofs of  $\mathbb{L}$  Log  $\mathbb{L}$  Criteria for Mean Behavior of Branching Processes. *The Annals of Probability*, 23(3), 1125-1138. <https://doi.org/10.1214/aop/1176988176>
- Lyons, R., & Peres, Y. (2016). *Probability on Trees and Networks*. Cambridge University Press.
- Marckert, J.-F., & Mokkadem, A. (2003). The Depth First Processes of Galton–Watson Trees Converge to the Same Brownian Excursion. *The Annals of Probability*, 31(3), 1655-1678. <https://doi.org/10.1214/aop/1055425793>
- Norris, J. R. (1998). *Markov Chains* (1.<sup>a</sup> ed.). Cambridge University Press.
- Pitman, J. (2006). *Combinatorial Stochastic Processes* (J. Picard, Ed.; Vol. 1875). Springer-Verlag. <https://doi.org/10.1007/b11601500>
- Revuz, D. (1984). *Markov Chains* (2.<sup>a</sup> ed.). North-Holland.
- Revuz, D., & Yor, M. (1999). *Continuous Martingales and Brownian Motion* (Vol. 293). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-06400-9>
- Sharpe, M. (1988). *General Theory of Markov Processes*. Academic Press.
- Whitt, W. (1980). Some Useful Functions for Functional Limit Theorems. *Mathematics of Operations Research*, 5(1), 67-85.