



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

POSGRADO EN CIENCIAS MATEMÁTICAS
FACULTAD DE CIENCIAS

EL INVARIANTE DEL ÁRBOL GENERADOR EN LOS
SISTEMAS DE CICLOS DE PRIMER RETORNO

T E S I S

QUE PARA OBTENER EL GRADO ACADÉMICO DE:
MAESTRO EN CIENCIAS

P R E S E N T A:

JOSÉ MIGUEL SALAZAR MONTIEL

DIRECTOR DE TESIS: DR. RICARDO GÓMEZ AÍZA

MÉXICO D. F.

NOVIEMBRE 2009



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

El Invariante del Árbol Generador
en Sistemas de Ciclos de Primer Retorno

José Miguel Salazar Montiel

Agradecimientos

Escribir los agradecimientos siempre se me ha hecho una tarea difícil, como agradecer (o culpar) a tantas personas por ser el individuo que soy en este momento. Por lo cual intentare agradecer a los que recuerde, de antemano pido perdón si hay alguna omisión de mi parte.

A mi Jefa, no hay palabras para agradecer todo lo que hace por mi. A mi Jefe por su consejo y paciencia. A mi Hermano y Cheche por ayudarme a crecer. Al monstruo por hacerme reír cuando lo necesito. A Diana por todo el cariño que me da. A a toda mi familia.

A Ricardo por sus enseñanzas paciencia y amistad durante tanto tiempo. Al Dr. Emilio por su sabiduria y por esas charlas tan interesantes. Maria Emilia por sus consejos he interés en este trabajo. Así como a Juancho por la dedicación a este trabajo y al Dino. A Hortensia por su contribución a que este sea un trabajo de calidad.

Agradezco al posgrado en Ciencias Matemáticas de la UNAM. En especial a su director Manuel Falconi por aceptar ser sinodal de este trabajo y todo el apoyo que se me ha brindado . Así como a Coco y Laurita y Tere que sin ellas es probable que no me hubiera podido inscribir ni a segundo semestre.

Asi como a mis compañeros Natalia, Julian, Gloria, Hiroki, Hector, Valente, Cesar, Maria etc.

También a todos aquellos que ha pesar de ellos acabe la maestría Pablito, Doña, Condor, Jose, Maria, Enano, Marti, Carlitos, Cesar, Queto, Itzel, Aranda , Picazo (por que yo si lo pongo), Betsa, Goyo, Manuel, Mau,, Gerry, Al Condy en general, Ayami, Chut, Pamela, Renata, Fa, Aramara, etc. Y a todos aquellos que no mencione GRACIAS.

Resumen

En [13] Lind y Tuncel introducen el concepto de *invariante del árbol generador* para cadenas de Markov con un número finito de estados. Su motivación se justifica cuando demuestran que el invariante del árbol generador es invariante bajo conjugaciones de cadenas de Markov. Su definición está basada en las características combinatorias de las gráficas que generan a los σ -espacios de tipo finito subyacentes (las cadenas de Markov *topológicas*, llamadas así por B. Parry [16]).

En esta Tesis extenderemos el concepto del invariante del árbol generador a cadenas de Markov con un número *numerable* de estados. En particular, vamos a considerar una clase apropiada de cadenas de Markov, a saber, aquellas que están representadas por las gráficas de los *sistemas de ciclos de primer retorno* definidos en [3, 9, 10]. Para estos sistemas estableceremos una forma explícita del valor de sus invariantes del árbol generador. De esta manera obtendremos una lista de valores, a saber, los invariantes de árbol generador de cada uno de los sistemas de ciclos de primer retorno, y nuestro objetivo central será el de determinar qué clase de información acerca del sistema se obtiene a partir de esta lista.

La motivación original para estudiar el invariante del árbol generador fue la siguiente. En [9] se introdujeron y clasificaron los *isomorfismos de palabra mágica*, una clase de isomorfismo entre cadenas de Markov (ver Definición 4.7). Una cadena de Markov (irreducible) siempre es isomorfa a sus sistemas de ciclos de primer retorno vía un isomorfismo de palabra mágica (Proposición 3.1 en [9]), por lo que la siguiente pregunta es natural: ¿El invariante del árbol generador es invariante bajo isomorfismos de palabra mágica? En esta Tesis demostramos que en general el invariante del árbol generador de los sistemas de ciclos de primer retorno de una cadena de Markov difieren del invariante del árbol generador del sistema original. Por lo tanto, en virtud de la pregunta original que nos planteamos, es natural cuestionarse cuándo coinciden los invariantes del árbol generador de los sistemas de ciclos de primer retorno. En esta Tesis damos respuesta a esta pregunta cuando el número de estados no excede a tres (ver Teoremas 4.11 y 4.13), a saber, demostraremos que para una matriz estocástica de tamaño a lo más tres, los invariantes del árbol generador de los sistemas de ciclos coinciden si y sólo si la matriz es doblemente estocástica. Más precisamente, demostraremos que para matrices estocásticas de tamaño a lo más tres, los invariantes del árbol generador de los sistemas de ciclos de primer retorno están dados por las razones del invariante del

árbol generador global con cada uno de los invariantes del árbol generador local, y veremos que éstas razones coinciden precisamente cuando la matriz es doblemente estocástica.

Cálculos numéricos muestran que éste es el caso para matrices (irreducibles) arbitrarias (el problema en el caso general radica en diagonalizar la matriz que resulta de remover de la matriz original el renglón y la columna correspondientes al vértice base del sistema de ciclos de primer retorno). Basándose en los resultados anteriores formulamos la *Conjetura de los Invariantes de los Árboles Generadores de Sistemas de Ciclos de Primer Retorno*, la cual demostramos en esta Tesis para los casos $n = 2, 3$.

Las matrices doblemente estocásticas han sido ampliamente estudiadas y existe una extensa literatura acerca de ellas. Entre los resultados más importantes destacan el hecho de que el conjunto de matrices doblemente estocásticas de tamaño n es un politopo en \mathbb{R}^{n^2} . Este politopo tiene como puntos extremos a las matrices de permutaciones, resultado obtenido por Birkhoff en 1946 (ver [2]). Los resultados que presentamos muestran las capacidades del árbol generador para poder medir ciertas simetrías en cadenas de Markov, a saber doble estocasticidad.

Índice general

1. Dinámica simbólica	1
1.1. σ -Espacios	1
1.1.1. σ -Espacios de tipo finito	1
1.1.2. Gráficas dirigidas y σ -espacios de tipo finito	5
1.1.3. Puntos periódicos y funciones zeta de Artin-Mazur	8
1.1.4. Entropía topológica	12
1.2. Homomorfismo de σ -Espacios	19
1.2.1. Homomorfismos en σ -espacios	19
1.2.2. σ -equivalencia fuerte	24
2. Cadenas de Markov	29
2.1. Cadenas de Markov	29
2.1.1. Cadenas de Markov	29
2.1.2. Clases de comunicación	35
2.1.3. Cadenas de Markov <i>numerables</i> y su clasificación	37

2.1.4. Teoría de Perron-Frobenius y distribuciones estacionarias	40
2.1.5. Entropía de teoría de la medida.	41
2.2. Matrices doblemente estocásticas	42
2.2.1. Permanentes	45
2.2.2. El Teorema de Birkhoff-Von Neumann	47
3. El invariante del árbol generador	49
3.1. El invariante del árbol generador	49
3.2. Invarianza del invariante del árbol generador	51
4. Sistemas de ciclos de primer retorno	57
4.1. Sistemas de ciclos de primer retorno	57
4.1.1. Presentación del problema	61
4.2. El invariante del árbol generador en los sistemas de ciclos de primer retorno	64
4.2.1. El “2-shift”	66
4.2.2. El “3-shift”	68
5. Conclusiones	83

Capítulo 1

Dinámica simbólica

En este capítulo presentamos una breve introducción de los aspectos más relevantes de dinámica simbólica que están relacionados con el trabajo que desarrollaremos a lo largo de la Tesis. Está dividido en dos secciones, comenzando con la sección §1.1 de σ -espacios seguida de la sección §1.2 de homomorfismos.

1.1. σ -Espacios

El propósito de esta sección es el de presentar brevemente las definiciones y resultados básicos de dinámica simbólica que usaremos. Esta sección está dividida en dos partes. En la subsección §1.1.1 describimos uno de los objetos centrales de la dinámica simbólica, los σ -espacios de tipo finito. En la sección §1.1.2 describimos cómo los σ -espacios de tipo finito están representados por gráficas dirigidas. No es la intención de esta sección la de presentar un panorama completo de la dinámica simbólica en general y tan sólo hemos incluido las definiciones y resultados relevantes a nuestro estudio. Para una referencia completa del tema el lector puede consultar [12].

1.1.1. σ -Espacios de tipo finito

Comenzaremos definiendo nuestros objetos de estudio, los σ -espacios. Sea \mathcal{A} un conjunto finito, el cual llamaremos *alfabeto*. Sea $\mathcal{A}^{\mathbb{Z}}$ el conjunto de todas las funciones $\mathbb{Z} \rightarrow \mathcal{A}$, es decir, $\mathcal{A}^{\mathbb{Z}}$ consiste de todas las sucesiones “bi-infinitas” de elementos del alfabeto. A $\mathcal{A}^{\mathbb{Z}}$ le llamaremos el *σ -espacio completo* o “full shift” en \mathcal{A} . Formalmente, tenemos la siguiente definición.

Definición 1.1. Sea \mathcal{A} un conjunto finito de símbolos. El σ -espacio completo en \mathcal{A} es el conjunto

$$\mathcal{A}^{\mathbb{Z}} = \{x = (x_i)_{i \in \mathbb{Z}} = \dots x_{-3}x_{-2}x_{-1}.x_0x_1x_2x_3\dots \mid x_i \in \mathcal{A} \quad \forall i \in \mathbb{Z}\}.$$

Dotamos a $\mathcal{A}^{\mathbb{Z}}$ con la topología producto, al considerar el conjunto de símbolos \mathcal{A} como un espacio topológico con la topología discreta. Este espacio es metrizable con la llamada métrica de Cantor.

Definición 1.2. Sean $x, y \in \mathcal{A}^{\mathbb{Z}}$ dos puntos cualesquiera ($x \neq y$). Definimos la métrica de Cantor como:

$$d_C(x, x) = 0 \quad \text{y} \quad d_C(x, y) = 2^{-K}$$

donde $K = \min\{i \geq 0 \mid x_i \neq y_i \text{ o } x_{-i} \neq y_{-i}\} < \infty$.

Es conveniente introducir cierta notación. Dados $n, m \in \mathbb{Z}$ con $n \leq m$, denotaremos por $[n, m]$ al intervalo de números enteros, siempre que en el contexto esté claro que no se está refiriendo al intervalo de números reales. De igual forma, $[n, m)$ denotará al conjunto $\{n, n+1, \dots, m-1\}$ y definiremos $(n, m]$ y (n, m) de forma similar. Así, dado $x \in X$, establecemos la notación

$$x_{[n, m]} = x_n \dots x_m.$$

y diremos que el bloque $x_{[n, m]}$ es de longitud $l(x_{[n, m]}) = m - n + 1$. De manera similar definiremos $x_{[n, m)}$, $x_{(n, m]}$ y $x_{(n, m)}$ (éste último está definido para $m > n$), y tendremos $l(x_{[n, m]}) = l(x_{(n, m]}) = m - n$ y $l(x_{(n, m)}) = m - n - 1$.

Una base de la topología inducida por la métrica de Cantor está formada por los conjuntos que se conocen como *cilindros*. A continuación se da la definición de éstos, pero la prueba de que son una base de los abiertos generados por la métrica de Cantor no se encuentra en este trabajo. Para la demostración véase [12] y [19].

Definición 1.3. Sea $w \in \mathcal{A}^n$ con $n \geq 1$. Sea $j \in \mathbb{Z}$. El cilindro generado por w en la coordenada j es el conjunto

$$[w; j] = \{x \in \mathcal{A}^{\mathbb{Z}} \mid x_{[j, j+n-1]} = w\}.$$

Tenemos definido ya el espacio básico en el cual se trabaja. Necesitamos definir ahora una “dinámica” en dicho espacio. Definimos la función σ como sigue.

Definición 1.4. Se define la función $\sigma: \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$ de un punto $x \in \mathcal{A}^{\mathbb{Z}}$ como $\sigma(x)_i = x_{i+1}$ para toda $i \in \mathbb{Z}$.

La función σ recorre los índices una posición hacia la izquierda y es sencillo verificar que es continua (recuérdese que los cilindros son una base de la topología y claramente la imagen inversa de un cilindro bajo la función σ es de nuevo un cilindro). Con la función σ podemos definir en general el concepto de σ -espacio (no necesariamente completo).

Definición 1.5. Un σ -espacio consiste de un subconjunto cerrado de sucesiones de símbolos $X \subset \mathcal{A}^{\mathbb{Z}}$ tal que $\sigma(X) = X$.

Es claro que el σ -espacio completo es un σ -espacio, sin embargo no todos los σ -espacios son el σ -espacio completo. Demos un ejemplo de lo anterior.

Ejemplo 1.6. Sea $\mathcal{A} = \{0, 1\}$ y sea X los puntos $x \in \mathcal{A}^{\mathbb{Z}}$ que no contienen al bloque 00. Es fácil ver que tal ejemplo es un σ -espacio y no es todo $\{0, 1\}^{\mathbb{Z}}$. Se conoce a X como el σ -espacio *aureo*.

Un σ -espacio X hereda la topología inducida por el σ -espacio completo. Esta topología también tiene como base a los conjuntos cilindros, ahora definidos para cada $w \in \mathcal{A}^n$ y $j \in \mathbb{Z}$ por

$$[w; j] = \{x \in X \mid x_{[j, j+n-1]} = w\} \quad (1.7)$$

y claramente esta topología es metrizable por la métrica de Cantor restringida.

Dentro de los σ -espacios, enfocaremos nuestra atención en una subclase particular, a saber los σ -espacios de tipo *finito* (el Ejemplo 1.6 es un instancia). Para poder definirlos usaremos el hecho de que un σ -espacio lo determina un conjunto de bloques prohibidos, es decir, que los bloques de dicho conjunto no se encuentran en ninguno de los puntos de nuestro σ -espacio. Enunciaremos este hecho formalmente. Sea

$$\mathcal{A}^* = \bigcup_{n=0}^{\infty} \mathcal{A}^n$$

(aquí \mathcal{A}^0 representa al bloque *vacío*).

Proposición 1.8 (Ver [12], Proposición 1.3.4.). *Sea X un σ -espacio sobre el alfabeto \mathcal{A} . Entonces existe $\mathcal{F} \subset \mathcal{A}^*$ tal que*

$$X = X_{\mathcal{F}} = \{x \in \mathcal{A}^{\mathbb{Z}} \mid \nexists w \in \mathcal{F} \text{ tal que } w = x_{[k, k+l(w)-1]} \text{ para alguna } k \in \mathbb{Z}\}.$$

En virtud de la proposición anterior, una lista de bloques prohibidos $\mathcal{F} \subset \mathcal{A}^*$ define un σ -espacio $X_{\mathcal{F}}$, a saber, el conjunto de puntos $x \in \mathcal{A}^{\mathbb{Z}}$ tales que para todo bloque $w \in \mathcal{F}$ de longitud n y para todo $k \in \mathbb{Z}$, $x_{[k, k+n-1]} \neq w$. Ahora bien, es sencillo observar que dos listas de palabras prohibidas distintas pueden generar el mismo σ -espacio. Por ejemplo, si utilizamos nuestro Ejemplo 1.6 con $\mathcal{F} = \{00\}$ y $\mathcal{F}' = \{00, 000, 0000, \dots\}$, en ambos casos el σ -espacio es el mismo, es decir, $X_{\mathcal{F}} = X_{\mathcal{F}'}$. Sin embargo, hay una única lista de bloques prohibidos maximal como veremos a continuación. Primero definimos el *lenguaje* de X como el conjunto $\mathcal{L}(X)$ formado por los bloques en \mathcal{A}^* que *sí* ocurren en algún punto de X , es decir,

$$\mathcal{L}(X) = \{w \in \mathcal{A}^* \mid \exists x \in X \text{ tal que } w = x_{[0, 0+l(w)-1]}\}$$

(se toma por convención considerar a los bloques que se encuentran en la coordenada 0, pues al ser un σ -espacio cualquier bloque w que aparece en $y_{[k, k+l(w)-1]}$ se encuentra en $\sigma^k(y)_{[0, 0+l(w)-1]}$). El conjunto maximal de bloques prohibidos es entonces

$$\mathcal{F}_X = \mathcal{A}^* \setminus \mathcal{L}(X).$$

Es sencillo observar que $X = X_{\mathcal{F}_X}$ y que si $\mathcal{F} \subset \mathcal{A}^*$ es tal que $X = X_{\mathcal{F}}$, entonces $\mathcal{F} \subset \mathcal{F}_X$.

Dado un σ -espacio X sobre un alfabeto \mathcal{A} , denotaremos por $\mathcal{B}_n(X)$ al conjunto de bloques permitidos en X de longitud n , es decir,

$$\mathcal{B}_n(X) = \mathcal{L}(X) \cap \mathcal{A}^n = \{w \in \mathcal{L}(X) \mid l(w) = n\}.$$

Ya que diversas listas de bloques prohibidos pueden determinar un mismo σ -espacio, a través de estas listas de bloques prohibidos de los σ -espacio podremos clasificarlos en dos clases fundamentales, aquellos σ -espacios que se les puede asociar una lista finita de bloques prohibidos que lo determinan y aquellos que no.

Definición 1.9. *Un σ -espacio es de tipo finito si existe una lista finita de bloques prohibidos que lo determina.*

Así, si X es un σ -espacio sobre un alfabeto finito \mathcal{A} , entonces X es de tipo finito si y sólo si existe $\mathcal{F} \subset \mathcal{A}^*$ tal que $X = X_{\mathcal{F}}$. En esta Tesis nos fijaremos en los σ -espacios de tipo finito y los denotaremos por *SFT* por sus siglas en inglés.

Sea X un σ -espacio sobre un alfabeto \mathcal{A} . Para poder determinar si un punto

$$x = \dots x_{-3}x_{-2}x_{-1}.x_0x_1x_2x_3\dots$$

está en nuestro σ -espacio X , tenemos que buscar en toda la sucesión de símbolos si es que existen bloques prohibidos dentro de ésta, es decir, si $\mathcal{F} \subset \mathcal{A}^*$ es tal que $X = X_{\mathcal{F}}$, entonces $x \in X$ si y sólo si para cada $w \in \mathcal{F}$,

$$x_{[i, i+l(w)-1]} \neq w \text{ para toda } i \in \mathbb{Z}.$$

Cuando X es un SFT, es posible construir un conjunto de bloques prohibidos de la misma longitud que determinan el mismo σ -espacio. Esto se logra como a continuación describimos. Como X es un SFT, existe un conjunto finito $\mathcal{F} \subset \mathcal{A}^*$ tal que $X = X_{\mathcal{F}}$. Tomamos el máximo de las longitudes de los bloques prohibidos en \mathcal{F} , y puesto que la lista es finita, dicha longitud es un número finito, digamos $M \geq 1$ (el caso $M = 1$ corresponde a los σ -espacios completos). Nuestro nuevo conjunto consistirá de todos los bloques de dicha longitud que contienen a los bloques prohibidos originales como sub-bloques. De manera precisa, definiremos $\mathcal{F}' \subset \mathcal{A}^M$ por

$$\mathcal{F}' = \{u_1 \dots u_M \in \mathcal{A}^M \mid \exists w \in \mathcal{F} \text{ tal que } w = u_i \dots u_{i+l(w)} \text{ para alguna } i = 1, \dots, M - l(w)\}.$$

Es sencillo ver que esta nueva lista genera el mismo σ -espacio, es decir, $X = X_{\mathcal{F}} = X_{\mathcal{F}'}$. Entonces, para todo σ -espacio de tipo finito, existe una lista de bloques prohibidos, donde todos los bloques son del mismo tamaño.

Definición 1.10. *Un σ -espacio de tipo finito tiene memoria M si puede ser descrito por una colección de bloques prohibidos todos de tamaño $M + 1$.*

Durante este trabajo se utilizarán gráficas dirigidas como representaciones de σ -espacios de tipo finito. Los σ -espacios que inducen las gráficas dirigidas son de memoria $M = 1$. Por lo tanto, estrictamente hablando, no todo SFT será un σ -espacio inducido por una gráfica dirigida. Sin embargo, esto siempre ocurre *módulo* una conjugación. Este es el tema de la siguiente sección.

1.1.2. Gráficas dirigidas y σ -espacios de tipo finito

Sea $G = (V, E)$ una gráfica dirigida. Primero será conveniente tener una notación básica para varios conceptos.

- $i(e)$ es el vértice inicial de la arista $e \in E$.
- $t(e)$ es el vértice final de la arista $e \in E$.
- E_I es el conjunto de aristas con vértice inicial $I \in V$.
- E^J es el conjunto de aristas con vértice final $J \in V$.
- $E_I^J = E_I \cap E^J$.

Existen dos formas en que una gráfica dirigida $G = (V, E)$ determina un σ -espacio de tipo finito, a saber, los σ -espacios por *vértices* y los σ -espacios por *aristas*. El σ -espacio por vértices es aquel cuyo alfabeto es el conjunto V y cuyo conjunto de bloques prohibidos es el conjunto $\mathcal{F} = \{(u, v) \in V \times V \mid (u, v) \notin E\}$; a este σ -espacio por vértices lo denotaremos por \hat{X}_G . El σ -espacio por aristas es aquel cuyo alfabeto es el conjunto E y el conjunto de bloques prohibidos es el conjunto $\mathcal{F} = \{(a, b) \in E \times E \mid t(a) \neq i(b)\}$; al σ -espacio generado de esta forma, lo denotaremos por X_G . Claramente, los σ -espacios tanto por vértices como por aristas son de tipo finito, siempre y cuando la gráfica G sea finita (es decir, tanto V como E son conjuntos finitos).

A cada SFT puede ser considerado tanto un SFT por vértices como por aristas a través de la *presentación en bloques* del σ -espacio (aquí, “considerado” querrá decir módulo una *conjugación*, ver Definición 1.43).

Definición 1.11. *Sea X un σ -espacio. y sea $N \geq 1$. Definimos un nuevo SFT*

Explícitamente, sea X un SFT de memoria M y sea \mathcal{F} un conjunto de bloques prohibidos de tamaño $M+1$ que determinan a X . Definimos una gráfica dirigida $G_M = (V, E)$, donde $V = \mathcal{B}_M(X)$ y para cada par de vértices $I = x_1x_2 \dots x_M \in V$ y $J = y_1y_2 \dots y_M \in V$, pondremos una arista de I a J si $x_2 \dots x_M = y_1 \dots y_{M-1}$.

Definición 1.12. Sea X un SFT de memoria M y sea $G_M = (V, E)$ la gráfica descrita en el párrafo anterior. Decimos que X_{G_M} es la presentación en M -bloques de X .

De esta forma, a lo largo de toda esta Tesis, vamos a identificar a los σ -espacios de tipo finito con las gráficas dirigidas que determinan las presentaciones en bloques (ya sea que se consideren como σ -espacios por vértices o por aristas).

Ejemplo 1.13. Pensaremos en el σ -espacio completo en un alfabeto de dos elementos como el σ -espacio por vértices que determina la gráfica que se muestra en la Figura 1.1.

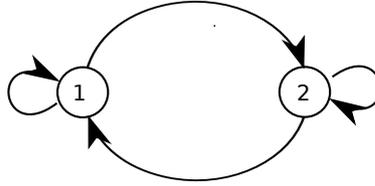


Figura 1.1: El 2-shift completo

Veamos algunos de los conceptos de la teoría de gráficas con los que estaremos trabajando.

Sea $G = (V, E)$ una gráfica, donde V es el conjunto de los vértices y E es el conjunto de aristas. Sea $|V| = n$ y enumeremos los vértices de G de 1 a n . Vamos a definir la *matriz de adyacencia* $A \in \mathcal{M}_{n \times n}(\mathbb{N})$ de G como sigue. La entrada $A_{i,j} = k$ donde $k \in \mathbb{N}$ es el número de flechas que van del vértice i al vértice j . Recordemos que un *camino* en G de longitud m es una sucesión de aristas (e_1, \dots, e_m) donde $t(e_i) = i(e_{i+1})$ para toda $i = 1, \dots, m-1$. La matriz de adyacencia tiene varias propiedades, entre las cuales es importante mencionar la siguiente.

Observación 1.14. El número de caminos de longitud $m \geq 1$ del vértice i al vértice j en G es $A_{i,j}^m$, donde A es la matriz de adyacencia de G .

Esta observación será *fundamental* para obtener expresiones explícitas de los sistemas de ciclos (ver ecuación (1.27)), la cual se utilizará en el Capítulo 4 en la demostración de nuestros resultados.

Demos un ejemplo de una gráfica sencilla, similar al 2-shift completo del ejemplo 1.13, en esta ocasión agregando una nueva arista de 1 a 2 (ver Figura 1.2).

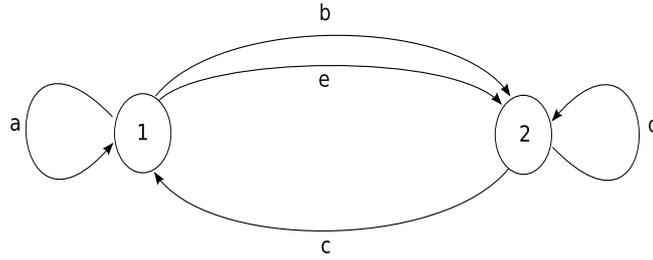


Figura 1.2: El 2-shift con una arista más

Ejemplo 1.15. Consideremos la gráfica de la Figura 1.2, cuya matriz de adyacencia es

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}.$$

Tenemos que

$$A^2 = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}.$$

Así pues, vemos que hay 3 caminos de longitud 2 del vértice 1 al vértice 1, a saber, (a, a) , (b, c) y (e, c) , hay 4 caminos de longitud 2 del vértice 1 al vértice 2, a saber, (a, b) , (a, e) , (b, d) y (e, d) , hay 2 caminos de longitud 2 del vértice 2 al vértice 1, a saber, (c, a) y (d, c) y finalmente hay 3 caminos de longitud 2 del vértice 2 al vértice 2, a saber, (c, b) , (c, e) y (d, d) .

Aunque la matriz A^n nos da el número de caminos de longitud n de un vértice a otro, no nos dice *qué* caminos son. Podría ser importante saber cuáles son estos caminos. Para esto, así como en [14], es conveniente considerar matrices sobre anillos más generales (referimos al lector a [11] para conceptos básicos de álgebra abstracta). Concretamente, pensemos en el anillo que resulta al considerar el grupo integral sobre el grupo libre generado por el conjunto de aristas. Recordemos que el grupo integral sobre un conjunto A se define como el conjunto de sumas formales

$$\mathbb{Z}[A] = \{a_1[x_1] + \cdots + a_n[x_n] \mid a_i \in \mathbb{Z} \text{ y } x_i \in A \text{ para toda } i = 1, \dots, n \text{ y } n \in \mathbb{N}\}$$

con la convención de que $a[x] + b[x] = (a + b)[x]$. Si el conjunto A tiene además una estructura de grupo, entonces definimos un producto inducido por la regla $(a[x])(b[y]) = ab[xy]$ y obtenemos así un anillo que será conmutativo si el grupo A lo es.

Si consideramos a \hat{E} como el grupo libre generado por el conjunto de aristas E , tendremos que la matriz sobre $\mathbb{Z}[\hat{E}]$ asociada a la gráfica dirigida de la Figura 1.2 es

$$\hat{A} = \begin{pmatrix} a & b + e \\ c & d \end{pmatrix}.$$

Tomamos ahora \hat{A}^2 ,

$$\hat{A}^2 = \begin{pmatrix} a^2 + bc + ec & ab + ae + bd + ed \\ ca + dc & cb + ce + d^2 \end{pmatrix}.$$

La entrada $\hat{A}_{1,1}^2$ son los posibles caminos del vértice 1 al vértice 1 de longitud 2. Tomemos a la entrada $\hat{A}_{1,1}^2 = a^2 + bc + ec$ lo que quiere decir que los caminos de longitud 2 del vértice 1 al vértice 1 es pasar 2 veces por a o pasar por b y por c o bien ir por e y regresar por c . De igual forma se puede hacer para todos los posibles caminos.

✦

Es conocido que el uso del grupo integral sobre grupos (multiplicativos) resulta de gran utilidad en el proceso de codificación que involucran los sistemas de ciclos de primer retorno (abordaremos un poco de esto posteriormente).

1.1.3. Puntos periódicos y funciones zeta de Artin-Mazur

Los puntos periódicos están asociados a las órbitas periódicas, y entender el comportamiento de las órbitas es el fundamental en cualquier dinámica. Por lo tanto, esto es también fundamental en el estudio de los σ -espacios. Este es el tema que abordaremos ahora puesto que están en estrecha relación con los sistemas de ciclos de primer retorno que veremos más adelante.

Definición 1.16. Sea X un SFT. Decimos que $x \in X$ es un punto periódico si existe $n \geq 1$ tal que

$$\sigma^n(x) = x$$

y en este caso decimos que x es de periodo n . Si $x \in X$ es un punto periódico, entonces decimos que x es de periodo mínimo $n \geq 1$ si $\sigma^n(x) = x$ y para todo $m < n$ se tiene que $\sigma^m(x) \neq x$. Definimos

$$p_n = |\{x \in X \mid x \text{ es periódico de periodo } n\}|$$

y

$$q_n = |\{x \in X \mid x \text{ es periódico de periodo mínimo } n\}|.$$

Definimos también el periodo de X como

$$\text{per}(X) = \text{mcd}\{p_n \mid n \geq 1\}.$$

X es aperiódico si $\text{per}(X) = 1$.

Claramente $q_n \leq p_n$ para toda $n \geq 1$. De hecho, como es de esperarse, la sucesión (p_n) de los puntos periódicos determina a la sucesión (q_n) de los puntos periódicos de periodo mínimo y viceversa, como veremos a continuación. Primero, como ya observamos, todo punto periódico de periodo mínimo n es un punto periódico de periodo n , es decir,

$$\{x \in X \mid x \text{ es periódico de periodo mínimo } n\} \subseteq \{x \in X \mid x \text{ es periódico de periodo } n\}.$$

Ahora bien, para cada $m \in \mathbb{N}$ tal que $m|n$, tenemos que

$$\{x \in X \mid x \text{ es periódico de periodo } m\} \subseteq \{x \in X \mid x \text{ es periódico de periodo } n\}$$

pues si $x \in X$ es de periodo m y $n = ma$ con $a \in \mathbb{N}$, entonces como $\sigma^m(x) = x$ tenemos que

$$\sigma^n(x) = \sigma^{am}(x) = \underbrace{\sigma^m(\sigma^m(\dots\sigma^m(x)))}_{a \text{ veces}} = x.$$

De ésta manera tenemos que la sucesión (p_n) puede obtenerse a partir de la sucesión (q_n) de la siguiente forma:

$$p_n = \sum_{m|n} q_m.$$

La fórmula explícita para la sucesión (q_n) en términos de la sucesión (p_n) es más complicada e involucra a la *función de inversión de Möbius*.

Definición 1.17. La función de inversión de Möbius se define para todo entero $n \geq 1$ por

$$\mu(n) = \begin{cases} 0 & \text{si existe un primo } p \text{ tal que } p^2|n \\ (-1)^k & \text{si } n = p_1 \dots p_k \text{ con } p_i \neq p_j \text{ primos distintos.} \end{cases}$$

Proposición 1.18 (Ver [12], pág. 191).

$$q_n = \sum_{k|n} \mu\left(\frac{n}{k}\right) p_k.$$

Cuando se tiene una sucesión de números es común codificarla en una *función generadora* (sobre funciones generadoras en general, referimos al lector a [8, 5]). En el caso de los puntos periódicos, la función generadora por excelencia es la función zeta de Artin-Mazur, la cual definimos a continuación.

Definición 1.19. Sea X un σ -espacio. La función zeta de Artin-Mazur de X (o simplemente la función zeta) es

$$\zeta_X(t) = \exp\left(\sum_{n=1}^{\infty} \frac{p_n}{n} t^n\right).$$

Ya que la función zeta está determinada por los puntos periódicos, también está determinada por los puntos periódicos de periodo mínimo, como vemos a continuación.

Proposición 1.20 (Ver [12], pág. 201). Sea X un σ -espacio. Entonces

$$\zeta_X(t) = \prod_{n=1}^{\infty} \frac{1}{1 - t^{q_n}}. \quad (1.21)$$

Ahora veremos que la función zeta de X es una función racional ya que

$$\zeta_X(t) = \frac{1}{\det(I - tA)}$$

donde A es la matriz de adyacencia de la gráfica de la presentación en M -bloques de X , donde M es la memoria de X . En virtud de lo dicho anteriormente tenemos la siguiente proposición:

Proposición 1.22. *Si tenemos un σ -espacio determinado a partir de una gráfica con matriz de adyacencia A de tamaño $r \times r$, con polinomio característico \mathcal{X}_A , entonces tenemos que*

$$\zeta_X(t) = \frac{1}{\text{tr } \mathcal{X}_A(t^{-1})} = \frac{1}{\det(I - tA)}.$$

PRUEBA. Denotemos las raíces de $\mathcal{X}_A(t)$ por $\lambda_1, \lambda_2, \dots, \lambda_r$ enlistadas por multiplicidad. Entonces el polinomio característico es de la siguiente forma:

$$\mathcal{X}_A(u) = (u - \lambda_1) \dots (u - \lambda_r) = u^r (1 - \lambda_1 u^{-1}) \dots (1 - \lambda_r u^{-1}) = u^r \det(I - u^{-1}A);$$

sustituyendo a u por t^{-1} se muestra que,

$$\mathcal{X}_A(t^{-1}) = t^{-r} (1 - \lambda_1 t) \dots (1 - \lambda_r t) = t^{-r} \det(I - tA).$$

o para nuestro caso

$$\frac{1}{\mathcal{X}_A(t^{-1})} = \frac{t^r}{(1 - \lambda_1 t) \dots (1 - \lambda_r t)} = \frac{t^r}{\det(I - tA)}$$

que es equivalente a

$$\frac{1}{\text{tr } \mathcal{X}_A(t^{-1})} = \frac{1}{(1 - \lambda_1 t) \dots (1 - \lambda_r t)} = \frac{1}{\det(I - tA)}. \quad (1.23)$$

Nos falta demostrar que

$$\zeta_X(t) = \frac{1}{\text{tr } \mathcal{X}_A(t^{-1})} = \frac{1}{\det(I - tA)}.$$

Sabemos que los puntos de periodo n en X están determinados biunívocamente por los bloques tamaño n que son posibles de concatenar con ellos mismos, es decir por los bloques $w \in \mathcal{B}_n(X)$ tales que $ww \in \mathcal{L}(X)$, y estos a su vez están determinados por la traza en la matriz A^n , puesto que en la traza se encuentran los ciclos de tamaño n . Es decir,

$$p_n(X) = \text{tr } A^n = \lambda_1^n + \lambda_2^n + \dots + \lambda_r^n.$$

Por lo que tenemos que

$$\begin{aligned}\zeta_X(t) &= \exp\left(\sum_{n=1}^{\infty} \frac{\lambda_1^n + \lambda_2^n + \cdots + \lambda_r^n}{n} t^n\right) \\ &= \exp\left(\sum_{n=1}^{\infty} \frac{\lambda_1^n t^n}{n} + \cdots + \sum_{n=1}^{\infty} \frac{\lambda_r^n t^n}{n}\right) \\ &= \exp\left(\sum_{n=1}^{\infty} \frac{\lambda_1^n t^n}{n}\right) \times \cdots \times \exp\left(\sum_{n=1}^{\infty} \frac{\lambda_r^n t^n}{n}\right) \\ &= \frac{1}{1 - \lambda_1 t} \times \cdots \times \frac{1}{1 - \lambda_r t}.\end{aligned}$$

La última igualdad se sigue de la conocida identidad

$$\ln \frac{x}{x-1} = \sum_{n=1}^{\infty} \frac{1}{nx^n}$$

y del hecho de que

$$\frac{x}{x-1} = \frac{1}{1-1/x}.$$

✦

Uno de los hechos fundamentales que ocuparemos en esta Tesis y que tiene que ver con cocientes de funciones zeta es el siguiente. Recordemos que si $G = (V, E)$ es una gráfica, entonces un *ciclo* en $v \in V$ es un camino cerrado que comienza y termina en v , y un *ciclo de primer retorno a $v \in V$* corresponde a un camino cerrado que cruza por v sólo al inicio y al final del camino. Denotaremos por

$$p_n(v) = \{\gamma \mid \gamma \text{ es un ciclo en } G \text{ en } v\}. \quad (1.24)$$

y

$$q_n(v) = \{\gamma \mid \gamma \text{ es un ciclo en } G \text{ de primer retorno a } v\}. \quad (1.25)$$

Proposición 1.26. *Sea $G = (V, E)$ una gráfica (simple) y sea \hat{X}_G el SFT por vértices. Sea*

$$f_v(t) = \sum_{n=1}^{\infty} |q_n(v)| t^n.$$

Entonces

$$f_v(t) = A_{vv}t + \sum_{n=0}^{\infty} \sum_{i,j \in V - \{v\}} A_{vi} (A^{(v)})_{ij}^n A_{jv} t^{n+2} \quad (1.27)$$

y también

$$f_v(t) = 1 - \frac{\zeta_A(t)}{\zeta_{A^{(v)}}(t)}$$

donde $A^{(v)}$ es la matriz que resulta de A al remover el renglón y la columna correspondientes a v .

Vamos a probar la ecuación (1.27) que es la que usaremos en las demostraciones de nuestros resultados más adelante.

PRUEBA. Claramente $q_1(v) = A_v v$ y esto justifica el término de grado uno en la ecuación (1.27). Si $n \geq 2$, entonces un ciclo de primer retorno a v consiste de la concatenación de una arista en E_v^i para alguna $i \in V - \{v\}$, seguida de un camino arbitrario de longitud $n \geq 0$ que no utiliza el vértice v y que termina en $j \in V - \{v\}$, seguido de una arista en E_j^v . En virtud de la Observación (1.14), la cantidad de caminos que resultan corresponde entonces al coeficiente del n -ésimo término de la serie de potencias $f_v(t)$. ✦

1.1.4. Entropía topológica

Sea X un σ -espacio. Si nos fijamos en el número $|\mathcal{B}_n(X)|$ de los posibles n -bloques que aparecen en X , nos da una idea de qué tan complejo es el σ -espacio X . Entre mayor (asintóticamente) sea el número de posibles n -bloques, más complicado el σ -espacio es. En lugar de utilizar la sucesión de $|\mathcal{B}_n(X)|$ para $n = 1, 2, \dots$, resumiremos el comportamiento de ésta calculando su tasa de crecimiento. Como se ve en [12], $|\mathcal{B}_n(X)|$ crece aproximadamente como 2^{cn} donde el valor c se le llama la constante de crecimiento.

Definición 1.28. Sea X un σ -espacio. La entropía (topológica) de X se define como

$$h(X) = \lim_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{B}_n(X)|.$$

Es de observar que si X es un σ -espacio con \mathcal{A} como alfabeto, tenemos que $|\mathcal{B}_n(X)| \leq |\mathcal{A}|^n$. Entonces

$$\frac{1}{n} \log |\mathcal{B}_n(X)| \leq \frac{1}{n} \log |\mathcal{A}|^n = \log |\mathcal{A}|$$

para toda n , por lo que tenemos que $h(X) \leq \log |\mathcal{A}|$. También si $X \neq \emptyset$, entonces $|\mathcal{B}_n(X)| \geq 1$ para toda n , así tenemos que $0 \leq h(X) < \infty$. Si $X = \emptyset$, entonces se tiene que $|\mathcal{B}_n(X)| = 0$ para toda $n \in \mathbb{N}$. Usando la definición 1.28 calculamos la entropía lo que nos queda como $h(X) = -\infty$.

Ejemplo 1.29. Determinemos la entropía de un σ -espacio clásico. Tomemos el σ -espacio X determinado por la siguiente matriz (éste es conocido como el σ -espacio *aureo*)

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Para poder utilizar la fórmula de la entropía, necesitamos conocer $|\mathcal{B}_m(X)|$. Como sabemos, el número de bloques de tamaño m está determinado por las entradas de la matriz A^m . Si diagonalizamos la matriz A podemos obtener con mayor facilidad dichas entradas. Por lo que

tomando el polinomio característico de A , $t^2 - t - 1$, obtenemos los valores característicos de A , es decir, las raíces $\lambda = \frac{1+\sqrt{5}}{2}$ y $\mu = \frac{1-\sqrt{5}}{2}$, con los que podremos diagonalizar la matriz.

De forma que para calcular a A^m tenemos que si

$$P = \begin{pmatrix} \lambda & \mu \\ 1 & 1 \end{pmatrix}$$

y

$$P^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -\mu \\ -1 & \lambda \end{pmatrix},$$

entonces

$$P^{-1}AP = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

Por lo que

$$A^m = P \begin{pmatrix} \lambda^m & 0 \\ 0 & \mu^m \end{pmatrix} P^{-1}.$$

Usando el hecho $\det A = \lambda\mu = -1$ tenemos que

$$A^m = \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda^{m+1} - \mu^{m+1} & -\lambda^{m+1}\mu + \mu^{m+1}\lambda \\ \lambda^m - \mu^m & -\lambda^m\mu + \mu^m\lambda \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda^{m+1} - \mu^{m+1} & \lambda^m - \mu^m \\ \lambda^m - \mu^m & \lambda^{m-1} - \mu^{m-1} \end{pmatrix}.$$

Si sustituimos a las entradas por una función de la forma $f_m = (\lambda^m - \mu^m)/\sqrt{5}$, tenemos que

$$A^m = \begin{pmatrix} f_{m+1} & f_m \\ f_m & f_{m-1} \end{pmatrix}.$$

Obsérvese que $\lambda^2 = \lambda + 1$, utilizando inducción se tiene que $\lambda^{m+2} = \lambda^{m+1} + \lambda^m$, de forma similar tenemos que $\mu^{m+2} = \mu^{m+1} + \mu^m$. Por lo que $f_{m+2} = f_{m+1} + f_m$ y entonces podemos calcular f_m como la sucesión de Fibonacci, es decir que el término m de la sucesión f_m está dado por la siguiente fórmula

$$f_m = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^m - \left(\frac{1-\sqrt{5}}{2} \right)^m \right] = \frac{1}{\sqrt{5}} [\lambda^m - \mu^m].$$

Calculemos $|B_m(X)|$:

$$|B_m(X)| = f_{m+1} + f_m + f_m + f_{m-1} = f_{m+2} + f_{m+1} = f_{m+3}.$$

Veamos que la entropía sólo depende de λ . Como $\mu^n/\lambda^n \rightarrow 0$ tenemos lo siguiente

$$\begin{aligned}
h(X) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{B}_n(X)| = \lim_{n \rightarrow \infty} \frac{1}{n} \log f_{n+3} \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\sqrt{5}} [\lambda^{n+3} - \mu^{n+3}] \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \left[\log \frac{1}{\sqrt{5}} + (n+3) \log \lambda + \log \left(1 - \frac{\mu^{n+3}}{\lambda^{n+3}} \right) \right] \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\sqrt{5}} + \lim_{n \rightarrow \infty} \frac{1}{n} (n+3) \log \lambda + \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(1 - \frac{\mu^{n+3}}{\lambda^{n+3}} \right) \\
&= \lim_{n \rightarrow \infty} \left(1 + \frac{3}{n} \right) \log \lambda = \log \lambda.
\end{aligned}$$

El σ -espacio en el Ejemplo 1.29 tiene por nombre el σ -espacio aureo puesto que tiene como entropía el logaritmo de la razón aurea $\lambda = \frac{1+\sqrt{5}}{2}$. ✦

Como hemos visto, se puede calcular (diagonalizando) la entropía de un σ -espacio que está determinado por una gráfica G con matriz de adyacencia $A \neq 0$ de tamaño $r \times r$. Esta técnica será una herramienta central en el desarrollo de los resultados que presentaremos en esta Tesis.

Para calcular la entropía de dicho σ -espacio necesitamos saber cuánto valen las entradas de A^n pues éstas determinan los posibles bloques de X . Es decir

$$|\mathcal{B}_n(X)| = \sum_{I, J=1}^r A_{I, J}^n.$$

Para calcular la entropía, es decir, encontrar la tasa de crecimiento de las entradas de A^n , el proceso de diagonalización no es siempre el más adecuado. Para las matrices que se ven en este trabajo, el teorema de Perron-Frobenius es la herramienta que nos ayuda a determinar el crecimiento de las entradas de la matriz.

Primero supongamos que tenemos un σ -espacio generado a partir de una gráfica fuertemente conexa con r vértices. Una gráfica $G = (V, E)$ es *fuertemente conexa* si para cualquier par de vértices $x, y \in V$, existe un camino dirigido que comienza en x y termina en y . Como mostramos anteriormente, para determinar la entropía necesitamos saber el valor de la entrada de la matriz A^n . Supongamos que la matriz tiene un vector característico positivo v , con valor característico λ . Para cada coordenada I del vector v se tiene que $Av_I > 0$ y $Av_I = \lambda v_I$. Aplicando A al vector v se obtiene $Av = \lambda v$ de igual forma $A^2v = \lambda^2v$, así sucesivamente tenemos que $A^n v = \lambda^n v$ con $n \geq 1$. Entonces para cada coordenada I del vector v tenemos que

$$\sum_{J=1}^r A_{I, J}^n v_J = \lambda^n v_I.$$

Sea $c = \min\{v_1, v_2, \dots, v_r\}$ y $d = \max\{v_1, v_2, \dots, v_r\}$. Entonces

$$c \sum_{J=1}^r A_{I,J}^n \leq \sum_{J=1}^r A_{I,J}^n v_J = \lambda v_I \leq d \lambda^n.$$

Dividiendo entre c y sumando sobre I obtenemos que

$$\sum_{I,J=1}^r A_{I,J}^n \leq \sum_{I=1}^r \frac{d}{c} \lambda^n = \left(\frac{rd}{c}\right) \lambda^n = d_0 \lambda^n$$

donde $d_0 = rd/c > 0$.

Veamos que también tiene una cota inferior, nótese que para cada I tenemos que

$$c \lambda^n \leq \lambda^n v_I = \sum_{J=1}^r A_{I,J}^n v_J \leq d \sum_{J=1}^r A_{I,J}^n \leq d \sum_{I,J=1}^r A_{I,J}^n$$

con $c_0 = c/d > 0$. Por lo que concluimos que

$$c_0 \lambda^n \leq \sum_{I,J}^r A_{I,J}^n.$$

Todo lo anterior lo podemos resumir en la siguiente proposición.

Proposición 1.30. *Sea $A \neq 0$ una matriz no negativa con vector característico positivo v . Entonces el valor característico correspondiente λ es positivo, y hay constantes c_0 y d_0 tales que*

$$c_0 \lambda^n \leq \sum_{I,J}^r A_{I,J}^n \leq d_0 \lambda^n.$$

Si A es la matriz de adyacencia de una gráfica que determina a un σ -espacio X_G , entonces $h(X_G) = \log \lambda$.

PRUEBA. En la discusión que precede al enunciado de la proposición se ha desarrollado ya gran parte de la prueba. Sólo falta demostrar la última afirmación que se obtiene simplemente al tomar el $\lim_{n \rightarrow \infty}$ en todas las desigualdades, es decir

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log(c_0 \lambda^n) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\sum_{I,J}^r A_{I,J}^n \right) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log d_0 \lambda^n.$$

Luego entonces

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log c_0 + \lim_{n \rightarrow \infty} \frac{1}{n} (n \log \lambda) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \left(\sum_{I,J}^r A_{I,J}^n \right) \leq \lim_{n \rightarrow \infty} \frac{1}{n} d_0 + \lim_{n \rightarrow \infty} \frac{1}{n} (n \log \lambda)$$

y así

$$\log \lambda \leq h(X) \leq \log \lambda.$$

✦

Como se ve, lo que necesitamos para poder encontrar las cotas es poder garantizar que exista un vector característico con entradas estrictamente positivas. En nuestro caso, el Teorema de Perron-Frobenius nos asegura la existencia de tal. Hasta ahora hemos considerado únicamente matrices con entradas enteras. El Teorema de Perron-Frobenius es verdadero en un casos más generales y será conveniente enunciarlo para matrices con entradas reales positivas.

Teorema 1.31 (Perron-Frobenius). *Sea $A \neq 0$ una matriz irreducible con entradas en \mathbb{R}^+ . Entonces A tiene un vector característico v_A con entradas estrictamente positivas y con valor característico $\lambda_A > 0$ tal que es geoméricamente y algebraicamente simple. Si μ es otro valor característico de A , entonces $|\mu| \leq \lambda_A$. Cualquier vector característico positivo para A es múltiplo de v_A .*

Cuando tenemos que A es una matriz real, irreducible y con entradas positivas, decimos que λ_A es el *valor Perron* de A , y a v_A se le llama el *vector Perron* de A . Obsérvese que el valor Perron es único y que el vector Perron es único salvo homotestas positivas. (Aquí no se da la demostración del teorema de Perron-Frobenius, pues es un resultado bastante conocido en el álgebra de matrices no-negativas. Para la prueba véase, e.g. [7]).

En caso de que A no sea irreducible, siempre es posible reescribir a A , módulo un reordenamiento del conjunto que la indexa, en forma diagonal por bloques, es decir, en la forma

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ * & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & A_k \end{pmatrix}$$

en donde cada A_i es irreducible. En este caso definimos el valor Perron de A como $\max\{\lambda_{A_i} \mid i = 1, \dots, k\}$.

Si tenemos un σ -espacio finito determinado por una gráfica G fuertemente conexa, entonces podemos afirmar lo siguiente:

Teorema 1.32. *Sea X un σ -espacio de tipo finito tal que la gráfica G que lo determina es fuertemente conexa y sea A su matriz de adyacencia. Entonces se cumple que $h(X) = \log \lambda_A$ donde λ_A es el valor Perron de A .*

PRUEBA. Por el teorema de Perron-Frobenius tenemos asegurada la existencia del valor λ en la Proposición 1.30 por lo que solo tenemos que aplicar dicha proposición.



El siguiente es un resultado clásico y lo enunciaremos porque será de utilidad más adelante.

Teorema 1.33 (Ver [12], Teorema 4.4.7). *Sea A una matriz irreducible y sea B una matriz con entradas no negativas, del mismo tamaño (finito) que A y tal que $0 \leq B \leq A$, es decir, $B_{ij} \leq A_{ij}$ para todas i, j . Supongamos que existen i, j tales que $B_{ij} < A_{ij}$. Entonces $\lambda_B < \lambda_A$.*

Entonces, si tenemos una gráfica finita, el σ -espacio que resultase de remover una arista o un vértice será de menor entropía. No es de sorprender entonces que la entropía también se relaciona fuertemente con los puntos periódicos y los valores p_n y q_n de la sección anterior por lo que a continuación veremos su relación a partir del siguiente teorema:

Teorema 1.34. *Sea X un σ -espacio definido a partir de un gráfica fuertemente conexa G . Entonces*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log p_n(X) = h(X).$$

PRUEBA. Sea X un sigma espacio definido a partir de una gráfica fuertemente conexa. Para demostrar la igualdad primero demostremos la siguiente desigualdad:

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log p_n(X) \leq h(X). \quad (1.35)$$

Para cada punto $x \in X$ de periodo n está unívocamente determinado por un bloque $x_{[0, n-1]} \in \mathcal{B}_n(X)$, por lo que $p_n \leq |\mathcal{B}_n(X)|$. Tomando logaritmo, dividiendo entre n y tomando el límite superior cuando $n \rightarrow \infty$, la igualdad se respeta.

Nos falta demostrar que se da la igualdad en nuestro caso, hagamos tal demostración. Como G es fuertemente conexa, entonces entre cualesquiera par de vértices, existe un camino de cierta longitud. Como la gráfica es finita, existe un N tal que para cualquier par de vértices I, J existe un camino de longitud igual o menor a N . Si $\pi \in \mathcal{B}_n(X)$, hay un camino τ de longitud menor o igual a N del vértice terminal $t(\pi)$ de π al vértice inicial $i(\pi)$ de π . Entonces $\pi\tau$ es un ciclo y $(\pi\tau)^\infty$ tiene periodo $n + |\tau|$, el cual esta entre n y $n + N$. De esta forma cada bloque en $\mathcal{B}_n(X)$ genera un punto en X con algún periodo entre n y $n + N$, bloques diferentes generan puntos diferentes. Por lo que tenemos que,

$$|\mathcal{B}_n(X)| \leq p_n(X) + p_{n+1}(X) + \cdots + p_{n+N}(X).$$

Para cada n hay un número entre los $p_n + k$ con $0 \leq k \leq N$ que es mayor a los otros. Digamos que tal número se da en $k = k(n)$. Entonces tenemos que,

$$|\mathcal{B}_n(X)| \leq (N + 1)p_{n+k(n)}(X).$$

Por lo que si tomamos la entropía tenemos,

$$\begin{aligned} h(X) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{B}_n| \\ &\leq \limsup_{n \rightarrow \infty} \left[\frac{1}{n} \log(N+1) + \log p_{n+k(n)}(X) \right] \\ &\leq \limsup_{m \rightarrow \infty} \frac{1}{m} \log p_m(X). \end{aligned}$$

con lo cual se prueba la otra desigualdad. ✦

De igual forma veamos como se relaciona la entropía con los puntos de periodo mínimo n . De igual forma que a partir del número de puntos periódicos de periodo n se puede calcular la entropía, vamos a calcular la entropía a partir del número de puntos periódicos de periodo mínimo n .

Por lo que tenemos la siguiente proposición:

Proposición 1.36. *Sea X un σ -espacio determinado a partir de una gráfica fuertemente conexa. Entonces se tiene que,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log q_n(X) = h(X).$$

PRUEBA. Como todo punto de periodo mínimo n es un punto periódico de periodo n , se sabe que $q_n(X) \leq p_n(X)$. Podemos utilizar la desigualdad (1.35), es decir,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log q_n(X) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log p_n(X) \leq h(X).$$

De esta forma al igual que en la prueba anterior sólo tenemos que demostrar la otra desigualdad. Si $h(X) = 0$ acabamos, por lo tanto supongamos que $h(X) > 0$. Como se demostró en la prueba anterior existe una sucesión de p_{n_i} tal que,

$$\limsup_{i \rightarrow \infty} \frac{1}{n_i} p_{n_i}(X).$$

Sea $\varepsilon > 0$. Entonces para alguna i lo suficientemente grande tenemos que

$$p_{n_i} > \exp(n_i h(X) - n_i \varepsilon). \quad (1.37)$$

También utilizando la demostración anterior tenemos que para una m lo suficientemente grande y toda $j \leq m$, tenemos

$$p_j < \exp(mh(X) + m\varepsilon). \quad (1.38)$$

Pero tenemos que

$$q_n(X) \geq p_n(X) - \sum_{j=1}^{n/2} P_j(X).$$

Entonces utilizando (1.37) y (1.38), para una i lo suficientemente grande, tenemos que

$$\begin{aligned}
q_{n_i} &\geq p_{n_i}(X) - \sum_{j=1}^{\frac{n_i}{2}} p_j(X) \\
&> \exp(n_i h(X) - n_i \varepsilon) - \left(\frac{n_i}{2}\right) \exp\left(\left(\frac{n_i}{2}\right) h(X) + \left(\frac{n_i}{2}\right) \varepsilon\right) \\
&= \exp(n_i h(X) - n_i \varepsilon) - \left(\exp(n_i h(X) - n_i \varepsilon)\right) \left(\left(\frac{n_i}{2}\right) \exp\left(-n_i \left(\frac{h(X)}{2} - \frac{3\varepsilon}{2}\right)\right)\right) \\
&= \left[\exp(n_i h(X) - n_i \varepsilon)\right] \left[1 - \left(\frac{n_i}{2}\right) \exp\left(-n_i \left(\frac{h(X)}{2} - \frac{3\varepsilon}{2}\right)\right)\right].
\end{aligned}$$

Tomando una ε lo suficientemente pequeña, tenemos que el multiplicando subrayado tiende a 1 cuando $i \rightarrow \infty$. Por lo que

$$\limsup_{i \rightarrow \infty} \frac{1}{n_i} \log q_{n_i}(X) \geq h(X).$$

Tomando la desigualdad anterior y la mostrada al principio de esta prueba tenemos la igualdad. ✦

Como hemos visto tanto el número de puntos periódicos, así como el número de puntos periódicos de periodo mínimo, determinan la entropía. Como éstos también determinan la función zeta, es claro que si dos σ -espacios tienen el mismo número de puntos periódicos p_n o el mismo número de puntos de periódicos de periodo mínimo q_n , entonces ambos tienen la misma función zeta y la misma entropía. Como veremos más adelante, esto nos sirve para determinar cierta “igualdad” entre los σ -espacios.

1.2. Homomorfismo de σ -Espacios

En esta sección presentamos definiciones y algunas propiedades de los homomorfismos entre los σ -espacios como son los códigos de bloques.

1.2.1. Homomorfismos en σ -espacios

En la dinámica simbólica necesitamos funciones entre la categoría de los σ -espacios. Para que estén bien definidas se requiere que respeten el hecho de ser σ -espacios, es decir que tienen que conmutar con la función σ . Comenzaremos con las funciones medibles con respecto a la σ -álgebra de Borel y que conmutan con σ . A este tipo de funciones las denominaremos *homomorfismos*.

Definición 1.39. *Un homomorfismo entre σ -espacios es una función medible (con respecto a la σ -álgebra de Borel) $\varphi: X \rightarrow Y$ tal que $\forall x \in X$ tenemos que $\varphi(\sigma(x)) = \sigma(\varphi(x))$, o en otras palabras,*

el siguiente diagrama conmute:

$$\begin{array}{ccc} X & \xrightarrow{\varphi} & Y \\ \sigma \downarrow & & \downarrow \sigma \\ X & \xrightarrow{\varphi} & Y \end{array}$$

Definición 1.40. Una función $\varphi: X \rightarrow Y$ es un isomorfismo si φ es un homomorfismo biyectivo y cuya inversa es un homomorfismo.

Entonces dos σ -espacios son isomorfos si existe un isomorfismo entre ellos.

Los isomorfismos entre SFT's fueron caracterizados por Ornstein y Friedmann en los años 70's (ver [15] y [6]). La caracterización es la siguiente:

Teorema 1.41 (Ornstein-Friedman). *Dos SFT's son isomorfos si y sólo si tienen la misma entropía y el mismo periodo.*

Los homomorfismos más frecuentes en la dinámica simbólica son las funciones continuas, las cuales están caracterizadas por el Teorema de Curtis-Lyndon-Hedlund (ver Teorema 1.44) como los códigos de bloques.

Definición 1.42. Sean X, Y σ -espacios. Una función $\varphi: X \rightarrow Y$ es un mapeo de bloques si existen enteros $m, a \geq 0$ y una función $\Phi: \mathcal{A}^{m+a+1} \rightarrow \mathcal{B}$ de forma que para cualesquiera $i \in \mathbb{Z}$ y $x \in X$, $\varphi(x)_i = \Phi(x_{[i-m, i+a]})$. Un mapeo de bloques está entonces representado de la siguiente forma:

$$\begin{array}{c} \dots x_{-i-m-1} \boxed{x_{-i-m} x_{-m+1} \dots x_{i+a-1} x_{i+a}} x_{i+a+1} \dots \\ \downarrow \Phi \\ \dots \varphi(x)_{i-1} \boxed{\varphi(x)_i} \varphi(x)_{i+1} \dots \end{array}$$

Decimos que φ tiene memoria m y anticipación a . A la función Φ se le llama la regla local de φ . Es claro que podemos encontrar una regla local con memoria y anticipación tan grande como queramos. En particular podemos suponer que $m = a = r$, en cuyo caso decimos que r es un radio de f .

Definición 1.43. Sean X y Y dos SFT's. Decimos que X y Y son conjugados si existe isomorfismo continuo $\varphi: X \rightarrow Y$ cuya inversa es también un isomorfismo continuo. En este caso decimos que φ es una conjugación.

Se han estudiado con amplitud este tipo de funciones aplicadas a los σ -espacios. Se conoce que mantienen diferentes propiedades a saber, puntos periódicos, equicontinuidad, puntos estables, transitividad, etc.

Los códigos de bloques son homomorfismos, pero cuándo una función que conmute con σ será un homomorfismo, o en otras palabras, ¿cuáles funciones son homomorfismos cuando tenemos la medida definida a partir de la métrica de Cantor? La respuesta a tal pregunta nos la da el teorema Curtis-Lyndon-Hedlund.

Teorema 1.44 (Curtis-Lyndon-Hedlund). Sea $\phi: X \rightarrow Y$ una función entre los σ -espacios X y Y . Entonces ϕ es un homomorfismo si y sólo si es un código de bloques.

Por falta de espacio en esta tesis la demostración de tal teorema no se encuentra en las páginas de este trabajo, pero se puede encontrar en [12] y [19]. Gracias al teorema anterior podemos afirmar que todas las funciones continuas bajo la métrica de Cantor que son homomorfismos, son códigos de bloques.

La siguiente pregunta resulta natural. ¿Si un código de bloques es biyectivo, su inversa siempre es un código de bloques?, es decir si tenemos un código de bloques biyectivo es de suponerse que este sea un isomorfismo, pero la respuesta es **no**, un código de bloques biyectivo no siempre tiene como función inversa a un código de bloques. Un código de bloques biyectivo cuya función inversa es un código de bloques se le llama *conjugación*. Dos σ -espacios de tipo finitos que sean isomorfos bajo códigos de bloques se dice que son *conjugados*. Se conocen varios invariantes bajo conjugación, siendo el más importante de todos la entropía [15, 6] (la demostración se encuentra más adelante en esta sección).

Se tienen clasificados los σ -espacios de tipo finito, módulo isomorfos que sean códigos de bloques, gracias al Teorema de Descomposición y al Teorema de la σ -Equivalencia Fuerte (ambos se encuentran en [12]). En vista de que la demostración de que el invariante del árbol generador es invariante de conjugación usa el Teorema de Descomposición, daremos su demostración pues nos ayudará a entender y visualizar mejor los procedimientos de las pruebas.

Para poder hacer la prueba será conveniente definir los *in-splittings* y *ex-splittings* para las gráficas. La definiciones formales se encuentran en [12] pero aquí daremos un breve recordatorio.

Definamos primero el in-splitting. Sea G una gráfica y k un vértice de G . Vamos a generar una nueva gráfica G' a partir de G . Sean E^k las flechas que tienen como vértice final a k , entonces

podemos tomar $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_m$ tal que $\bigcup_{i=1}^m \mathcal{F}_i = E^k$ y $\mathcal{F}_r \cap \mathcal{F}_s = \emptyset$ si $r \neq s$, una partición de las aristas con vértice final k . Entonces en G' tenemos un vértice k_i por cada conjunto \mathcal{F}_i y los restantes vértices son los mismos que en G .

Las aristas de G son las mismas si no tienen como vértice inicial o final al vértice k . Si tenemos una arista con vértice inicial k y vértice final x en G , añadimos una arista en G' con vértice inicial k_i y vértice final x , para cada $i \in \{1, 2, \dots, m\}$. Si tenemos una arista con vértice final k y vértice inicial x entonces $k \in \mathcal{F}_i$ para algún $i \in \{1, 2, \dots, m\}$, colocamos una arista con vértice inicial x y vértice final k_i en G' .

De forma análoga se define el out-splitting. Dada una gráfica G y k un vértice tomamos una partición de las aristas que tienen como vértice final a k , $E_k = \bigcup_{i=1}^m \mathcal{F}_i$ donde $\mathcal{F}_i \cap \mathcal{F}_j \neq \emptyset$ si $i \neq j$. Vamos a generar una nueva gráfica G'' donde por cada \mathcal{F}_i vamos a tomar un nuevo vértice k_i y los demás vértices de G los dejamos igual.

Las adyacencias de los vértices que sean distintos a k se mantienen igual. Por cada arista e que va de un vértice w al vértice k en G tenemos una arista de w a cada uno de los nuevos vértices k_i . Para cada arista con vértice inicial k se encuentra contenida en algún \mathcal{F}_i tal arista tiene como vértice inicial a k_i en G'' y como vértice final al original en G , en caso de que éste sea k se tiene que añadir una arista hacia todos los vértices que representan a k en G'' . Para cada k_i tenemos que $E_{k_i} = \mathcal{F}_i$.

Teorema 1.45 (Teorema de Descomposición). *Cualquier conjugación de un σ -espacio por aristas a otro, es la composición de los códigos splittings y amalgamations.*

PRUEBA. Gracias a la presentación en bloques de un σ -espacio podemos suponer que cualquier código de bloques es un 1-bloque. Para demostrar el Teorema de Descomposición necesitamos utilizar el siguiente lema y aplicarlo varias veces.

Lema 1. *Sean G, H gráficas, X_G denota al σ -espacio generado por las aristas de G . Sea $\phi: X_G \rightarrow X_H$ una conjugación cuya inversa tiene memoria $m \geq 0$ y anticipación $n \geq 1$. Entonces hay out-splittings \tilde{G} de G y \tilde{H} de H tales que el diagrama*

$$\begin{array}{ccc} X_G & \xrightarrow{\psi_{G\tilde{G}}} & X_{\tilde{G}} \\ \phi \downarrow & & \downarrow \tilde{\phi} \\ X_H & \xleftarrow{\alpha_{\tilde{H}H}} & X_{\tilde{H}} \end{array}$$

conmuta, donde $\alpha_{\tilde{H}H}$ y $\psi_{G\tilde{G}}$ son los códigos de bloques que resultan de los out-splittings. Además $\tilde{\phi}^{-1}$ tiene memoria m y anticipación $n - 1$.

PRUEBA. Sea \tilde{H} el out-splitting completo de H , entonces las aristas de \tilde{H} son de la forma h^k donde $h, k \in E(H)$ y $t(h) = i(k)$ en H . Sea $\phi = \Phi_\infty$ donde $\Phi: E(G) \rightarrow E(H)$. Para cada vértice $w \in V(G)$

daremos una partici3n de sus aristas E_w , definida de acuerdo a su imagen bajo Φ . Por lo que para cada $x \in V(G)$ y $h \in E(H)$ se define $E_w^h = \{g \in E_w \mid \Phi(g) = h\}$. Sea \tilde{G} la gr3fica que resulta de hacer el out-splitting bajo las particiones definidas anteriormente. Definamos $\tilde{\Phi}: E(\tilde{G}) \rightarrow E(\tilde{H})$ como $\tilde{\Phi}(g^h) = \Phi(g)^h$. De esta forma inducimos un c3digo de 1-bloque dado por $\tilde{\phi} = \tilde{\Phi}_\infty: X_{\tilde{G}} \rightarrow X_{\tilde{H}}$. El siguiente diagrama muestra la acci3n de estos c3digos.

$$\begin{array}{ccc}
 \dots g_{-3}g_{-2}g_{-1}.g_0g_1g_2g_3 \dots & \xrightarrow{\psi_{G\tilde{G}}} & \dots g_{-3}^{h_{-2}}g_{-2}^{h_{-1}}g_{-1}^{h_0}.g_0^{h_1}g_1^{h_2}g_2^{h_3}g_3^{h_4} \dots \\
 \downarrow \phi = \tilde{\Phi} & & \downarrow \tilde{\phi} = \tilde{\Phi}_\infty \\
 \dots h_{-3}h_{-2}h_{-1}.h_0h_1h_2h_3 \dots & \xleftarrow{\alpha_{\tilde{H}H}} & \dots h_{-3}^{h_{-2}}h_{-2}^{h_{-1}}h_{-1}^{h_0}.h_0^{h_1}h_1^{h_2}h_2^{h_3}h_3^{h_4} \dots
 \end{array}$$

Para demostrar que $\tilde{\phi}^{-1}$ tiene la memoria y anticipaci3n requerida, debemos ver que para cualquier $\tilde{y} \in X_{\tilde{H}}$, las coordenadas $[-m, n - 1]$ de \tilde{y} , determinan a la coordenada 0 de $\tilde{x} = \tilde{\phi}^{-1}(\tilde{y})$. Para esto escribimos $y = \alpha_{\tilde{H}H}(\tilde{y})$, $x = \phi^{-1}(y)$ y observamos que

$$\tilde{x}_0 = x_0^{y_1}.$$

Como $\tilde{y}_{[-m, n-1]}$ determina $y_{[-m, n]}$ que determina la coordenada x_0 se sigue que $\tilde{y}_{[-m, n-1]}$ determinan la coordenada \tilde{x}_0 . ✦

A continuaci3n aplicamos el Lema 1 repetidas veces para obtener la prueba del Teorema de Descomposici3n.

Como se mencion3 anteriormente en un principio, gracias a la presentaci3n de bloques de un σ -espacio, podemos suponer que el c3digo de bloques es 1-bloque. Sean m y n la memoria y la anticipaci3n de ϕ^{-1} . Podemos suponer que $m \geq 0$ y $n \geq 0$. Si $m = n = 0$, entonces ϕ es un c3digo de bloques de 1-bloque con inversa 1-bloque, entonces es solo una retiquetaci3n y terminamos. Supongamos que $n \geq 1$ por el Lema 1 hay una sucesi3n de out-splitings ψ_j y conjugaciones de 1-bloques $\tilde{\phi}$, donde $1 \leq j \leq n$, para las cuales se tiene que $\tilde{\phi}_j^{-1}$ tiene memoria m y anticipaci3n $n - j$, como se muestra en el siguiente diagrama.

$$\begin{array}{ccccccc}
 X_G & \xrightarrow{\psi_1} & X_{G_1} & \xrightarrow{\psi_2} & X_{G_2} & \longrightarrow \dots & \xrightarrow{\psi_n} & X_{G_n} \\
 \downarrow \phi & & \downarrow \tilde{\phi}_1 & & \downarrow \tilde{\phi}_2 & & & \downarrow \tilde{\phi}_n \\
 X_H & \xleftarrow{\alpha_1} & X_{H_1} & \xleftarrow{\alpha_2} & X_{H_2} & \longleftarrow \dots & \xleftarrow{\alpha_n} & X_{H_n}
 \end{array}$$

La funci3n inversa del c3digo de bloques $\tilde{\phi}_n$ tiene memoria m y anticipaci3n 0.

El Lema 1 puede ser usado para reducir la memoria del código de bloque, aplicando lo a la gráfica transpuesta (la cual resulta de transponer la matriz de adyacencia de la gráfica original o invertir las flechas de la gráfica original). Entonces tendremos una sucesión de in-splittings ψ_{n+k} y conjugaciones $\tilde{\phi}_{n+k}$ de 1-bloque para $1 \leq k \leq m$ tal que $\tilde{\phi}_{n+k}^{-1}$ tiene memoria $m - k$ y anticipación 0 como se muestra a continuación

$$\begin{array}{ccccccc} X_{G_n} & \xrightarrow{\psi_{n+1}} & X_{G_{n+1}} & \xrightarrow{\psi_{n+2}} & X_{G_{n+2}} & \longrightarrow & \cdots & \xrightarrow{\psi_{n+m}} & X_{G_{n+m}} \\ \downarrow \tilde{\phi}_n & & \downarrow \tilde{\phi}_{n+1} & & \downarrow \tilde{\phi}_{n+2} & & & & \downarrow \tilde{\phi}_{n+m} \\ X_{H_n} & \xleftarrow{\alpha_{n+1}} & X_{H_{n+1}} & \xleftarrow{\alpha_{n+2}} & X_{H_{n+2}} & \xleftarrow{\cdots} & & \xleftarrow{\alpha_{n+m}} & X_{H_{n+m}} \end{array}$$

En particular $\tilde{\phi}_{n+m}^{-1}$ tiene memoria 0 y anticipación 0, entonces es una retiquetación. Si ponemos estos diagramas juntos obtenemos el siguiente diagrama.

$$\begin{array}{cccccccccccccccc} X_G & \xrightarrow{\psi_1} & X_{G_1} & \xrightarrow{\psi_2} & X_{G_2} & \longrightarrow & \cdots & \xrightarrow{\psi_n} & X_{G_n} & \xrightarrow{\psi_{n+1}} & X_{G_{n+1}} & \xrightarrow{\psi_{n+2}} & X_{G_{n+2}} & \longrightarrow & \cdots & \xrightarrow{\psi_{n+m}} & X_{G_{n+m}} \\ \downarrow \phi & & \downarrow \tilde{\phi}_1 & & \downarrow \tilde{\phi}_2 & & & & \downarrow \tilde{\phi}_n & & \downarrow \tilde{\phi}_{n+1} & & \downarrow \tilde{\phi}_{n+2} & & & & \downarrow \tilde{\phi}_{n+m} \\ X_H & \xleftarrow{\alpha_1} & X_{H_1} & \xleftarrow{\alpha_2} & X_{H_2} & \xleftarrow{\cdots} & & \xleftarrow{\alpha_n} & X_{H_n} & \xleftarrow{\alpha_{n+1}} & X_{H_{n+1}} & \xleftarrow{\alpha_{n+2}} & X_{H_{n+2}} & \xleftarrow{\cdots} & & \xleftarrow{\alpha_{n+m}} & X_{H_{n+m}} \end{array}$$

Recorriendo el diagrama primero hacia la derecha, hacia abajo y a la izquierda podemos tener las sucesión de in-splittings y out-splittings para ir de una gráfica a otra. Por lo que ϕ es una composición de éstos.

Como colorario del teorema de descomposición tenemos el siguiente colorario

Corolario 1.46. Sean G y H gráficas esenciales. El σ -espacio por aristas X_G y X_H son conjugados si sólo si G se obtiene a partir de H a través de una sucesión de operaciones in-splitting, out-splitting, in-amalgamation y out-amalgamation.

1.2.2. σ -equivalencia fuerte

En esta sección enunciamos una condición entre SFT's que es equivalente a la conjugación. Esta es la caracterización más importante en términos de matrices y ha dado lugar a una teoría general llamada *Teoría de Williams*.

Sean $A, B \in \mathcal{M}(\mathbb{N})$ matrices cuadradas, si tomamos las gráficas que determinan dichas matrices a saber que estas son las matrices de adyacencia de dichas gráficas, y tomamos los SFT que determinan las gráficas, entonces ¿Cuándo los SFT son conjugados? El teorema de descomposición nos dice que si podemos llegar de la gráfica que determina la matriz A a la gráfica que determina la

matriz B a través de operaciones de splittings y amalgamations, entonces los SFT son conjugados. Para determinar si las matrices A y B determinan SFT conjugados sin conocer la sucesión de las operaciones en las gráficas tenemos la siguiente relación de equivalencia.

Definición 1.47. Sean A y B matrices cuadradas con entradas enteras no negativas. Una equivalencia elemental de A a B es un par de matrices (R, S) rectangulares con entradas enteras no negativas que satisfacen

$$A = RS \quad \text{y} \quad B = SR.$$

Y denotamos a este caso $A \cong_{R,S} B$.

Una “strong shift equivalence” o una *equivalencia fuerte* de longitud l de A a B es una sucesión de l equivalencias elementales

$$A = A_0 \cong_{R_1, S_1} A_1, A_1 \cong_{R_2, S_2} A_2 \dots A_{l-1} \cong_{R_l, S_l} A_l = B.$$

En este caso escribimos que $A \approx B$. Decimos que A es fuertemente equivalente a B si existe una equivalencia fuerte de A a B para alguna longitud. Es de observarse que las matrices A y B no necesariamente son del mismo tamaño.

Es fácil ver que ser *fuertemente equivalente* es una relación de equivalencia en el conjunto de matrices cuadradas de cualquier tamaño finito, ver que esta relación se da cuando los σ -espacios son conjugados se debe al trabajo de R.F. Williams en [21]. Nuestra intención será demostrar que si dos matrices son fuertemente equivalente, entonces los σ -espacios asociados a éstas, son conjugados. Para probar esto primero veremos que la equivalencia elemental puede ser usada para obtener una conjugación.

Sean A, B matrices cuadradas con entradas enteras no negativas, tales que hay una equivalencia elemental de A a B con las matrices R, S , es decir $A \cong_{R,S} B$. Contruyamos una gráfica $G_{R,S}$ a partir de las gráficas G_A y G_B que son las gráficas que determinan las matrices A y B respectivamente. Tomemos la unión ajena de las gráficas de G_A y G_B . Las aristas en G_A las llamaremos A -aristas, y las de G_B B -aristas. Para cada vértice I en G_A y cada vértice J en G_B le añadimos el número de aristas en la entrada $R_{I,J}$ con dirección de I a J , a estas la llamaremos R -aristas y de igual forma añadimos $S_{J,I}$ aristas de J a I y los llamamos S -aristas, completando la construcción de $G_{R,S}$. Por un R, S -camino entenderemos a una R -arista seguida de una S -arista formando así un camino en $G_{R,S}$. De forma similar definimos S, G -camino.

Sean I, I' vértices en G_A . Como $A = RS$ hay una biyección entre las aristas de I a I' y los R, S -caminos de I a I' . Podemos fijar una biyección y a cada A -arista le corresponderá un par de aristas, donde una es R -arista y la otra es S -arista, denotaremos $a \leftrightarrow r(a)s(a)$ o $a = a(sr)$ tal asignación. En forma similar tenemos que como $B = SR$, entonces para cualesquiera par de estados J, J' en G_B tenemos una biyección entre las aristas de J a J' y los S, R -caminos de J a J' . Fijamos una biyección y denotamos por $b \leftrightarrow S(b)r(b)$ o $b = b(sr)$ tal elección.

Podemos definir un código de bloques de 2-bloques, dando primero la regla local de $\Gamma_{R,S}: \mathcal{B}_2(X_A) \rightarrow \mathcal{B}_1(X_B)$ usando la biyección elegida de la siguiente forma:

$$\Gamma_{R,S}(a_0, a_1) = b(s(a_0)r(a_1)).$$

Lo que ocurre en realidad, es que un A -camino a_0a_1 le corresponde a un doble camino R, S -camino $r(a_0)s(a_0)r(a_1)s(a_1)$. La parte de en medio $s(a_0)r(a_1)$, es un S, R -camino que corresponde a una B -arista $b = b(s(a_0)r(a_1))$, el valor de $\Gamma_{R,S}$ en el bloque a_0a_1 . Fijando la elección de biyecciones definimos el código de bloques de la siguiente forma:

$$\gamma_{R,S} = (\Gamma_{R,S})_{\infty}^{[0,1]}.$$

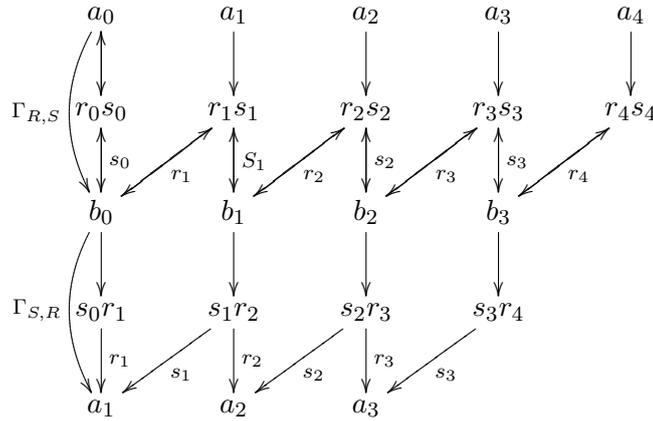
Si A y B son matrices de solo 0 – 1, entonces sólo hay una posible elección de las biyecciones.

Proposición 1.48. *Si la misma biyección es usada para definir $\gamma_{R,S}: X_A \rightarrow X_B$ y $\gamma_{S,R}: X_B \rightarrow X_A$, entonces*

$$\gamma_{S,R} \circ \gamma_{R,S} = \sigma_A \quad \text{y} \quad \gamma_{R,S} \circ \gamma_{S,R} = \sigma_B.$$

En particular los códigos de bloques $\gamma_{R,S}$ y $\gamma_{S,R}$ son conjugaciones.

PRUEBA. De las definiciones se ve que si $a_1a_2a_3 \in \mathcal{B}_3(X_A)$, entonces $\Gamma_{R,S}(a_1a_2)$ sigue a $\Gamma_{R,S}(a_0a_1)$ en G_B , por lo que $\gamma_{R,S}(X_A) \subset X_B$. De forma análoga $\gamma_{S,R}(X_B) \subset X_A$. Para visualizar ésto tenemos el siguiente diagrama. Denotemos por $r_i = (a_i)$, $s_i = (a_i)$.



El argumento expresado en términos de nuestra notación de biyecciones es de la siguiente forma:

$$\Gamma_{S,R} \circ \Gamma_{R,S}(a_0a_1a_2) = \Gamma_{S,R}(b(s(a_0)r(a_1))b(s(a_1)r(a_2))) = a(r(a_1)s(a_1)) = a_1.$$

Notese que de la segunda igualdad es valida por que usamos la misma biyección para definir $\Gamma_{S,R}$ y $\Gamma_{R,S}$, por lo que $r(b(s(a_0)r(a_1))) = a_1$, etc.

Un argumento similar muestra que $\gamma_{R,S} \circ \gamma_{S,R} = \sigma_B$. Como σ_A y σ_B son conjugaciones, se sigue que $\gamma_{R,S}$ y $\gamma_{S,R}$ también son conjugaciones. ✦

Utilizando la proposición anterior vemos que la equivalencia elemental entre matrices con entradas enteras no negativas, induce una conjugación, entre los σ -espacios que están determinados por las matrices. Si tomamos una equivalencia fuerte entre dos matrices A y B tenemos que gracias a que la composición de un número finito de conjugaciones, es una conjugación, entonces tenemos que los σ -espacios X_A y X_B determinados por las matrices son conjugados. A esta afirmación se le conoce como el teorema de clasificación y se enuncia de la siguiente manera.

Teorema 1.49 (Teorema de Clasificación). *Los σ -espacios X_A y X_B son conjugados si y sólo si las matrices A y B son fuertemente equivalentes.*

La demostracion de tal es aplicar lo visto anteriormente.

Para poder determinar si dos σ -espacios son conjugados, necesitamos encontrar el código de bloques que determina tal conjugación. Como ésto puede ser complicado (es de hecho indescidible), podemos encontrar invariantes bajo las conjugaciones y así determinar si los σ -espacios son conjugados o no. A continuación demostraremos que la función zeta, así como la entropía son invariantes bajo conjugación.

Para demostrar que tanto la función zeta y la entropía son invariantes bajo la conjugación, lo único que necesitamos demostrar para éste hecho, es ver que el número de puntos periódicos de un σ -espacios son invariantes bajo conjugación. Como la función zeta y la entropía están determinados por el número de puntos periódicos es claro que ambos tendrán la misma función zeta y la misma entropía. Enunciemos lo escrito anteriormente.

Teorema 1.50. *Si X y Y son σ -espacios conjugados, entonces la función zeta de X es igual a la función zeta de Y a saber que,*

$$\zeta_X(t) = \zeta_Y(t).$$

Así como también se da la igualdad con la entropía, es decir que $h(X) = h(Y)$.

El Teorema 1.50 se demuestra probando la siguiente proposición:

Proposición 1.51. *Sean X y Y dos sigma espacios, sea $\phi: X \rightarrow Y$ un código de bloques entre X y Y . Si $x \in X$ tiene periodo n en X , entonces $\phi(x)$ tiene periodo n en Y , y el periodo mínimo de $\phi(x)$ divide el periodo mínimo de x . Las conjugaciones respetan los puntos periódicos de periodo mínimo.*

PRUEBA. Demostremos la proposición anterior. Si $x \in X$ tiene periodo n , entonces $\sigma_X^n(x) = x$. Tenemos que,

$$\sigma_Y^n(\phi(x)) = \phi(\sigma_X^n(x)) = \phi(x).$$

por lo que $\phi(x)$ tiene periodo n . Si x tiene periodo mínimo n , entonces $\phi(x)$ tiene periodo n . Como el periodo de un punto siempre es múltiplo de su periodo mínimo. Por lo que tenemos que, si m es el periodo mínimo de $\phi(x)$, entonces m divide a n . Si ϕ es una conjugación, entonces m divide a n y n divide a m por lo tanto $m = n$. ✦

PRUEBA DEL TEOREMA 1.50. Utilizando la Proposición 1.51 tenemos que $q_n(X) = q_n(Y)$. Como las q_n 's determinan a las p_n 's tenemos que $p_n(X) = p_n(Y)$. Por la definición de la función zeta y la definición de la entropía (ver definición 1.28), tenemos que $\zeta_X(t) = \zeta_Y(t)$ y $h(X) = h(Y)$. ✦

De esta manera hemos definido dos invariantes bajo conjugación, a saber, la entropía y la función zeta. Es común que no sea posible encontrar el código de bloques entre dos espacios, por lo que para determinar si son conjugados una forma “fácil” de hacerlo es encontrando su entropía o su función zeta. Es decir, si dos σ -espacios tienen distinta entropía o distinta función zeta, entonces no son conjugados.

Entre más restricciones tenga un isomorfismo, más invariantes asociados tendrá. Entre los distintos invariantes que se conocen se encuentran los puntos periódicos, la función zeta, el grupo de la dimensión (ver [12]), entre otros. En particular se encuentra el invariante de árbol generador, el cual es un tema central de esta Tesis. Este invariante está definido en un contexto más general en el cual están involucradas medidas de probabilidad. Este es el tema del siguiente capítulo.

Capítulo 2

Cadenas de Markov

2.1. Cadenas de Markov

Ahora nos interesa definir medidas en los σ -espacios. Por convención, los σ -espacios serán vistos como espacios medibles al considerar la σ -álgebra de Borel. Queremos también que las medidas sean σ -invariantes, es decir, que si X es un σ -espacio, μ es una medida en X y $A \subset X$ es medible, entonces $\mu(A) = \mu(\sigma(A))$. En virtud del Teorema de Extensión de Caratheodory (ver, e.g. [17]), si definimos una medida en los cilindros del σ -espacio, dicha medida se extiende a todo el σ -espacio (ver sección §2.1.1 de este capítulo). Los espacios de medida resultantes son los que llamamos *Cadenas de Markov*. En este capítulo describiremos el proceso de construir esta clase de medidas sobre los σ -espacios. Asumimos cierta familiaridad con conceptos de teoría de probabilidad y de teoría de la medida. Como referencia el lector podrá consultar, por ejemplo, [4] y [8].

2.1.1. Cadenas de Markov

Definición 2.1 (CADENAS DE MARKOV COMO PROCESOS ESTOCÁSTICOS). *Sea (Ω, F, \mathbb{P}) un espacio de probabilidad, donde Ω es un conjunto finito o numerable, F es la σ -álgebra y \mathbb{P} es la medida de probabilidad. \mathbf{E} es un conjunto no vacío, finito o numerable. Una sucesión de variables aleatorias*

$$\{X_n: \Omega \rightarrow \mathbf{E}, \quad n = 0, 1, 2 \dots\}$$

se llama cadena de Markov con espacio de estados \mathbf{E} si satisface la condición de Markov, esto es, si para todo $n \geq 1$ y toda sucesión $x_0, x_1, \dots, x_{n-2}, x, y \in \mathbf{E}$ se cumple

$$\mathbb{P}(X_n = y \mid X_{n-1} = x, \dots, X_0 = x_0) = \mathbb{P}(X_n = y \mid X_{n-1} = x) \quad (2.2)$$

siempre que los eventos con los que se está condicionando tenga probabilidad positiva, es decir, siempre que

$$\mathbb{P}(X_{n-1} = x, \dots, X_0 = x_0) > 0.$$

La distribución de X_0 se llama distribución inicial y se denota con $\pi = (\pi_x)_{x \in \mathbf{E}}$ de forma que $\mathbb{P}(X_0 = x) = \pi_x \geq 0$ y $\sum_{x \in \mathbf{E}} \pi_x = 1$, es decir, π es un vector estocástico. Decimos que la cadena de Markov es homogénea si

$$\mathbb{P}(X_n = y \mid X_{n-1} = x) = \mathbb{P}(X_1 = y \mid X_0 = x) \quad \forall n = 1, 2, \dots$$

La definición anterior es la más general de una cadena de Markov (a tiempo discreto), y es la que comúnmente se asume en teoría de probabilidad (por ejemplo, ver [4, 8]). En este trabajo sólo consideraremos cadenas de Markov homogéneas. Dada una cadena de Markov homogénea $\{X_n\}_{n \in \mathbb{N}}$ sobre el conjunto de estados (finito) \mathbf{E} , podemos construir una gráfica dirigida $G = (V, E)$ de la siguiente manera. El conjunto de vértices V será el conjunto de estados \mathbf{E} y para cualesquiera $x, y \in \mathbf{E}$, pondremos una arista $e \in E$ con vértice inicial x y vértice terminal y si y sólo si $\mathbb{P}(X_1 = y \mid X_0 = x) > 0$ (la probabilidad de transición de x a y).

De lo anterior concluimos que dada una cadena de Markov homogénea, hay un σ -espacio subyacente (el σ -espacio por vértices que determina la gráfica dirigida descrita en el párrafo anterior), y por esto B. Parry llamo a los SFT's cadenas de Markov intrínsecas o “Markov shifts”. Veamos algunas formas equivalentes de definir cadenas de Markov y que serán de utilidad.

Definición 2.3 (CADENAS DE MARKOV SOBRE UN CONJUNTO FINITO). *Sea S un conjunto finito. Una cadena de Markov finita μ en S es una asignación de probabilidades iniciales $\mu(I) \geq 0$ para $I \in S$ y probabilidades condicionales $\mu(J|I) \geq 0$ para $I, J \in S$ tal que*

$$\sum_{I \in S} \mu(I) = 1$$

y

$$\sum_{J \in S} \mu(J|I) = 1 \quad \text{para toda } I \in S.$$

La idea es que S representa el conjunto de resultados, también llamados estados de un experimento (obsérvese que la asignación de probabilidades iniciales $\mu(I)$ en la Definición 2.3 corresponden a la distribución inicial π en la Definición 2.1). La probabilidad de que una sucesión de resultados I_0, I_1, \dots, I_n suceda, está definida por

$$\mu(I_0, I_1, \dots, I_n) = \mu(I_0)\mu(I_1 \mid I_0)\mu(I_2 \mid I_1) \dots \mu(I_n \mid I_{n-1}). \quad (2.4)$$

Dada una sucesión de resultados I_0, I_1, \dots, I_{n-1} que ya se han dado, la probabilidad condicional de que el siguiente resultado sea en particular I_n es

$$\mu(I_n \mid I_0, \dots, I_{n-1}) = \frac{\mu(I_0, I_1, \dots, I_n)}{\mu(I_0, I_1, \dots, I_{n-1})} = \mu(I_n \mid I_{n-1}) \quad (2.5)$$

(la primera igualdad se sigue de la definición de probabilidad condicional, a saber, $\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$), de forma que el siguiente resultado sólo depende del resultado más reciente, de forma que tenemos una propiedad de Markov (comparar las ecuaciones 2.5 y 2.2).

Ahora, usando la definición de cadenas de Markov sobre conjuntos finitos, definiremos cadenas de Markov en SFT's definidos por gráficas dirigidas. Hay dos posibilidades de escoger al conjunto finito de estados, a saber, a los vértices o a las aristas. Veamos cada una de éstas y consideremos algunos ejemplos para evaluar sus propiedades.

Definición 2.6 (CADENA DE MARKOV EN SFT'S POR ARISTAS). *Una cadena de Markov en un SFT por aristas definido por una gráfica $G = (V, E)$ es una asignación de probabilidades $\mu(I) \geq 0$ para $I \in V$ y una probabilidad condicional $\mu(e|i(e)) \geq 0$ para $e \in E$ tal que*

$$\sum_{I \in V} \mu(I) = 1$$

y

$$\sum_{e \in E_I} \mu(e|I) = 1 \quad \text{para toda } I \in V.$$

Para cada bloque $w = e_1 \dots e_n \in \mathcal{B}_n(X_G)$ y $k \in \mathbb{Z}$ definimos la probabilidad del cilindro $[w; k]$ por

$$\mu([w; k]) = \mu(i(e_1))\mu(e_1|i(e_1))\mu(e_2|i(e_2)) \dots \mu(e_n|i(e_n)). \quad (2.7)$$

Así como tenemos en gráficas la matriz de adyacencia, en el caso de las cadenas de Markov tenemos una matriz similar llamada matriz de probabilidades de transición, la cual definimos a continuación para las cadenas de Markov en SFT's por aristas.

Definición 2.8 (MATRIZ DE TRANSICIÓN DE CADENAS DE MARKOV EN SFT'S POR ARISTAS). *Tomemos una cadena de Markov sobre un SFT por aristas definido por la gráfica $G = (V, E)$. Definimos la matriz de probabilidades de transición o simplemente matriz de transición como una matriz $P \in \mathcal{M}_{|E| \times |E|}([0, 1])$, donde la entrada (e, f) representa la probabilidad de ir del estado (arista) e al estado (arista) f , es decir,*

$$P_{ef} = \mu(f | t(e)).$$

Tal matriz P es *estocástica*; es decir, la suma de los entradas en cada renglón es 1 (si las entradas en cada columna también sumaran uno, entonces sería *doblemente estocástica*). Veamos un ejemplo.

Ejemplo 2.9. Sea G la gráfica en la Figura 2.1 y supongamos que la probabilidad de estar en cualesquiera de los dos vértices 1 ó 2 es la misma, es decir que $\mu(1) = 1/2 = \mu(2)$.

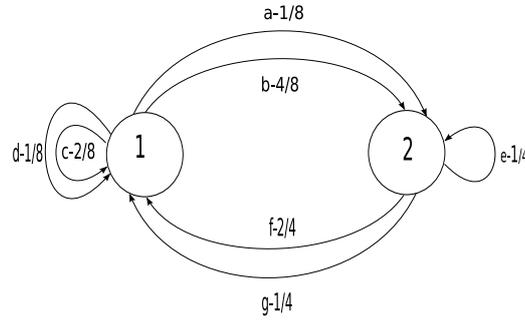


Figura 2.1: Ejemplo de una cadena de Markov en un SFT por aristas.

La matriz de probabilidades de transición es

$$P = \begin{pmatrix} \frac{2}{8} & \frac{1}{8} & \frac{4}{8} & \frac{1}{8} & 0 & 0 & 0 \\ \frac{1}{8} & \frac{1}{8} & \frac{4}{8} & \frac{1}{8} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2}{4} & \frac{2}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & \frac{2}{4} & \frac{2}{4} & \frac{1}{4} \\ \frac{2}{8} & \frac{1}{8} & \frac{4}{8} & \frac{1}{8} & 0 & 0 & 0 \\ \frac{1}{8} & \frac{1}{8} & \frac{4}{8} & \frac{1}{8} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2}{4} & \frac{2}{4} & \frac{1}{4} \end{pmatrix} \quad (2.10)$$

con los renglones y las columnas ordenados lexicográficamente. Ahora, por ejemplo, la probabilidad de tener el camino bg se obtiene de la siguiente forma:

$$\mu(bg) = \frac{1}{2} \times \frac{4}{8} \times \frac{1}{4} = \frac{4}{64}.$$

•

Será conveniente definir también cadenas de Markov en SFT's por vértices. Para esto, establezcamos la siguiente notación. Dada una matriz cuadrada A sobre \mathbb{R}_+ , definimos $A^\#$ como una matriz del mismo tamaño que A y definida por la regla

$$A_{ij}^\# = \begin{cases} 1 & \text{si } A_{ij} \neq 0 \\ 0 & \text{si } A_{ij} = 0 \end{cases} \quad (2.11)$$

y diremos que A y $A^\#$ son *compatibles*.

Ejemplo 2.9 (continuación). Consideremos la matriz P dada por (2.10). Tenemos entonces que

$$P^\# = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

de forma que $P^\#$ es la *matriz de incidencia* de la digráfica que determina a P (ver Figura 2.1.1).

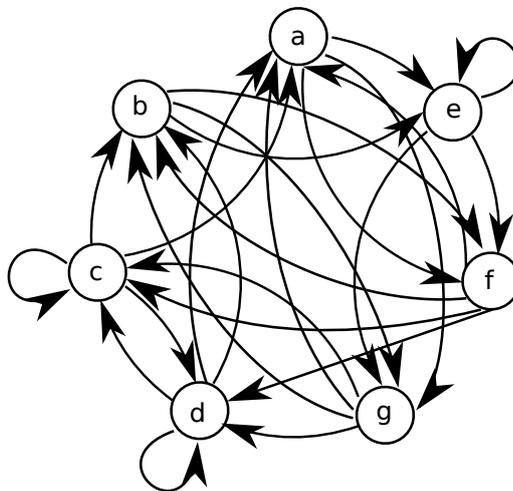


Figura 2.2: Representación como un SFT por vértices de un SFT por aristas.

En el anterior se ilustra cómo es posible pasar de un SFT por aristas a un SFT por vértices.

Definición 2.12 (CADENAS DE MARKOV EN SFT'S POR VÉRTICES). *Una cadena de Markov en un SFT por vértices dado por una gráfica (simple) $G = (V, E)$ es una asignación de probabilidades iniciales $\mu(I) \geq 0$ para cada $I \in V$ y probabilidad de transición $\mu(J|I) \geq 0$ para cada $I, J \in V$ de forma que*

$$\sum_{I \in V} \mu(I) = 1$$

y

$$\sum_{J \in V} \mu(J|I) = 1 \quad \text{para toda } I \in V.$$

Para cada bloque $w = v_1 \dots v_n \in \mathcal{B}_n(\hat{X}_G)$ y $k \in \mathbb{Z}$ definimos la probabilidad del cilindro $[w; k]$ por

$$\mu([w; k]) = \mu(v_1)\mu(v_2|v_1) \dots \mu(v_n|v_{n-1}). \quad (2.13)$$

Definición 2.14 (MATRIZ DE TRANSICIÓN DE CADENAS DE MARKOV EN SFT'S POR VÉRTICES). Tomemos una cadena de Markov finita en un SFT por vértices dado por una gráfica (simple) $G = (V, E)$. Definimos la matriz de probabilidades de transición o simplemente matriz de transición como una matriz $P \in \mathcal{M}_{|V| \times |V|}([0, 1])$, donde la entrada (I, J) representa la probabilidad de cruzar la arista del vértice I al J (si no existe, entonces el valor de la entrada es 0). De forma que

$$P_{IJ} = \mu(J | I).$$

Como se ilustró en los ejemplos, cualquier camino en nuestra digráfica, ya sea que se trate de una representación de un SFT por aristas o por vértices, es posible definir una medida de probabilidad que claramente es σ -invariante. En efecto, primero asignamos una medida a los cilindros con las ecuaciones (2.7) o (2.13) y observamos que no dependen del valor entero $k \in \mathbb{Z}$, lo que implica la σ -invarianza de la medida resultante y que se obtiene del Teorema de Extensión de Caratheodory. Las medida de un cilindro la vamos a interpretar entonces como la probabilidad condicional de un cierto camino w dado $i(w) = u$, el vértice inicial del camino w , por la probabilidad de “estar” en el vértice u .

En la práctica es más común el uso de los SFT's por vértices que el de por aristas, pues en general es más sencillo pensar en los estados de un proceso estocástico de Markov como los nodos de una gráfica simple (i.e. posiblemente con “loops” pero sin aristas múltiples). De esta manera es que se presentan en las aplicaciones: cadenas de Markov en SFT's por vértices descritas con matrices estocásticas indexadas por los vértices de una gráfica simple. En este caso, una de estas matrices estocásticas P induce una gráfica dirigida $G_P = (V, E)$ con $V = \{1, \dots, |P|\}$, donde $|P|$ denota el tamaño de P , y para cada $i, j \in V$, ponemos una arista de i a j si y sólo si $P_{i,j} > 0$, de forma que la matriz de adyacencia de G es precisamente $P^\#$, la matriz compatible con P definida anteriormente en (2.11). Entonces la matriz P , junto con un vector de probabilidades iniciales $\{\mu_I\}_{I \in V}$, induce una cadena de Markov en G donde la probabilidad condicional de $e \in E$ está dada por $\mu(e | i(e)) = P_{i(e), t(e)}$. De forma que si $w = x_0, \dots, x_n$ es un camino en la digráfica y $k \in \mathbb{Z}$, entonces la medida del cilindro $[w; k]$ es

$$\mu([w; k]) = \mu(x_0)P_{x_0, x_1} \dots P_{x_{n-1}, x_n}$$

(esta definición no depende del parámetro k , por lo que en efecto determina una cadena de Markov homogénea). Como es sabido los cilindros son la base de la topología inducida por la métrica de Cantor, por lo que se extiende una medida a todo el espacio de manera σ -invariante. Por la probabilidad de un camino w entenderemos

$$\mathbb{P}(w) = P_{x_0, x_1} \dots P_{x_{n-1}, x_n}. \quad (2.15)$$

Más adelante, cuando veamos Teoría de Perron-Frobenius en la sección §2.1.4, estableceremos a la *distribución estacionaria* como el vector de probabilidades iniciales *de facto*.

2.1.2. Clases de comunicación

A continuación definiremos algunos conceptos de las cadenas de Markov que son comunes en teoría de probabilidad. Observaremos algunas de estos conceptos desde varios puntos de vista tanto de teoría de gráficas como de dinámica simbólica.

Definición 2.16. *Supongamos que tenemos una cadena de Markov en un conjunto de estados S y matriz de transición $(P_{x,y})_{x,y \in S}$, para $x, y \in S$ se dice que:*

1. *De x se accede a y si existe $n \geq 0$ tal que $P_{x,y}^n > 0$ y se denota por $(x \rightarrow y)$.*
2. *x y y se comunican entre sí y se denota por $(x \leftrightarrow y)$ si se cumple que $(x \rightarrow y)$ y $(y \rightarrow x)$.*

$(x \leftrightarrow y)$ da lugar a una relación de equivalencia (véase [4] o [8]).

En la teoría de gráficas la definición anterior de acceder equivale a que existe un camino de longitud n entre los vértices x y y .

Definición 2.17. *Supongamos que tenemos una cadena de Markov (homogénea) con conjunto de estados S .*

1. *Una clase de equivalencia respecto a la relación (\leftrightarrow) se llama clase de comunicación.*
2. *Dado $x \in S$ su clase de comunicación se denota como*

$$C(x) = \{y \in S \mid x \leftrightarrow y\}$$

3. *Se dice que un subconjunto de estados $C \subset S$ es cerrado si ningún estado de $S \setminus C$ puede ser accedido desde un estado en C , es decir que $P_{x,y}^m = 0$ para todo $x \in C$, todo $y \in S \setminus C$ y todo $m \geq 1$.*

El primer concepto se refiere a las componentes fuertemente conexas de una digráfica. El tercero hace referencia a una subgráfica C , que no tiene aristas de salida hacia el resto de la gráfica.

Los conceptos anteriores nos ayudan a determinar cuándo es posible que a partir de un estado $x \in S$ se pueden visitar a los otros estados en S , es decir, si es posible que una sucesión de eventos que en algún momento visita al estado $x \in S$ pueda visitar a cualquier otro estado $y \in S$. Aplicando la clase de equivalencia a la gráfica que determina a un SFT, podemos dividir la gráfica en clases de equivalencia o *clases de comunicación*, en donde los estados de cada clase se comunican entre si. Estas clases de equivalencia determinan a σ -subespacios dentro del σ -espacio

original. Como las propiedades que nos interesan en esta Tesis se pueden estudiar a partir de estos σ -subespacios, enfocaremos nuestra atención en las cadenas de Markov que cumplan que entre cualesquiera estados se comunican. Por lo cual a continuación definiremos un concepto utilizando los conceptos anteriores.

Definición 2.18. *Se dice que una cadena de Markov (homogénea) con conjunto de estados S es irreducible si cumple con cualquiera de las siguientes condiciones (equivalentes):*

(A) *Desde cualquier estado de S se puede acceder a cualquier otro.*

(B) *Todos los estados se comunican entre sí.*

(C) *$C(x) = S$ para algún $x \in S$.*

(D) *$C(x) = S$ para todo $x \in S$.*

(E) *El único conjunto cerrado (ver 3 en la Definición 2.17) es el total.*

La demostración de que las condiciones son equivalentes se encuentra en [4] o [8] y no se presentan dentro de esta Tesis. A continuación se encuentra la definición de una matriz *irreducible* en el contexto de la dinámica simbólica. Sólo demostraremos algunas de las equivalencias de que las condiciones anteriores son equivalentes a la siguiente definición.

Definición 2.19. *Sea P la matriz de transición de una cadena de Markov. P es irreducible si para cualesquiera par de estados i, j existe $n \in \mathbb{N}$ tal que $P_{i,j}^n > 0$. Decimos que P mezcla si además podemos escoger una n independientemente de los estados i, j , es decir, que existe $n \in \mathbb{N}$ tal que $P_{i,j}^n > 0$ para todo i, j .*

Proposición 2.20. *Sea P una matriz estocástica. Cada una de las condiciones (A), (B), (C), (D) y (E) es equivalente a que P sea irreducible.*

PRUEBA. Es suficiente demostrar que cualquiera de las condiciones en la definición 2.18 es equivalente a la definición 2.19, como ejercicio se hacen algunas más.

(A) \Leftrightarrow Def 2.19 . Si desde cualquier estado se puede acceder a cualquier otro, por la definición 2.16, es claro que esto ocurre si y sólo si existe $n \in \mathbb{N}$ tal que $P_{i,j}^n > 0$.

(B) \Leftrightarrow Def 2.19 . Claramente (A) \Leftrightarrow (B) y por lo tanto (B) es equivalente a la definición 2.19.

(C) \Rightarrow Def 2.19 . Si $C(x) = S$ para algún $x \in S$, entonces quiere decir que desde el estado x se puede acceder a cualquier estado $y \in S$. Entonces existe $m(y) \in \mathbb{N}$ tal que $P_{x,y}^m > 0$. Como (\leftrightarrow) es una relación de equivalencia, de igual forma existe $k(y) \in \mathbb{N}$ tal que $P_{y,x}^k > 0$. Utilizando otra

vez que (\leftrightarrow) es una relación de equivalencia para cualquier par de estados $y, y' \in S$ tenemos que $(y \leftrightarrow x)$ y $(x \leftrightarrow y')$ por lo que $(y \leftrightarrow y')$ entonces existe $n \in \mathbb{N}$ tal que $P_{y,x}^n > 0$.

(D) \Rightarrow Def 2.19 . Esta demostración es análoga a la anterior.

(E) \Rightarrow Def 2.19 . Como el único conjunto cerrado es el total, entonces no hay estado al cual no se pueda acceder desde cualquier estado de S , por lo que tenemos que la clase de comunicación de cualquier estado es el total S y estamos otra vez como en el inciso (C)

•

La definición anterior se relaciona con un importante concepto en las gráficas, llamado *conexidad*. Recordemos que una gráfica $G = (V, E)$ es *fuertemente conexa* si para cualesquiera par de vértices $x, y \in V$, existe un camino dirigido que comienza en x y termina en y . Se puede ver que para cualquier par de vértices existe un $n \in \mathbb{N}$ tal que hay un camino de longitud n entre este par de vértices, que es la definición de irreducibilidad aplicado a la matriz de adyacencia de la gráfica. Recordemos que si tomamos la matriz de adyacencia A de la gráfica del sistema que determina la matriz P , tenemos que $A_{i,j}^n$ es el número de caminos dirigidos de longitud n entre el vértice i y el vértice j . Si la gráfica es fuertemente conexa, entonces dicho camino existe y $P_{i,j}^n > 0$ que es la suma de las probabilidades condicionales de los caminos de longitud n del vértice i al vértice j , dado el vértice i .

2.1.3. Cadenas de Markov *numerables* y su clasificación

En el capítulo 4 vamos a trabajar con σ -espacios más generales puesto que vamos a considerar gráficas infinitas (numerables). Por esta razón es conveniente especificar bien lo que entenderemos por muchos de los conceptos que hemos presentado hasta entonces, pero ahora en el contexto de conjuntos *numerables* de estados, así como también observar la clasificación de estados que se origina en base a las propiedades de recurrencia de las cadenas. Primero, se asume que el conjunto de estados es ahora un conjunto numerable con la topología discreta. Entonces, por ejemplo, los σ -espacios ahora cesan de ser compactos en general. Es importante distinguir entonces a esta nueva clase de σ -espacios de los SFT's que consideramos anteriormente.

Definición 2.21. *Un Markov shift es el σ -espacio que resulta al considerar un σ -espacio definido por una gráfica dirigida (no necesariamente finita).*

Entonces los SFT's son los Markov shift's que corresponden a gráficas finitas.

La pérdida de la compacidad es sólo una de muchas propiedades que pueden cambiar en este nuevo contexto de espacios de estados numerables, y estos cambios dependerán de las probabilidades de transición. Aunque muchos de los conceptos se extienden de manera trivial, en esta sección vamos

hacer explícitos algunos de éstos para así establecer las condiciones necesarias para desarrollar el trabajo subsecuente.

Dadas dos matrices A y B reales indexadas por un conjunto numerable \mathcal{S} , el *producto* AB se define como la matriz indexada por \mathcal{S} y que está determinada por

$$(AB)_{uv} = \sum_{i \in \mathcal{S}} A_{ui} B_{iv}$$

con $u, v \in \mathcal{S}$. Este producto es una generalización natural del clásico producto de matrices finitas. De esta forma, si $n \geq 0$, A^n tiene sentido como el producto de A consigo mismo un total de n -veces (definimos $A^0 = Id$, donde Id es la matriz identidad de tamaño igual al tamaño de A). Decimos que A es con entradas no negativas es *irreducible* si siempre que $i, j \in \mathbb{N}$, existe un entero $N = N_{i,j} > 0$ tal que $P_{ij}^N > 0$. *Únicamente vamos a considerar matrices irreducibles (y por lo tanto no negativas), a menos de que se indique lo contrario.*

Sea P una matriz estocástica indexada por los números naturales $\mathbb{N} = \{1, 2, \dots\}$, de forma que para toda $i \in \mathbb{N}$,

$$\sum_{j=1}^{\infty} P_{ij} = 1.$$

A una matriz estocástica P infinita también se le tiene asociada su matriz equivalente $P^\#$ dada por (2.11), y podemos pensar en esta última como la matriz de adyacencia de una gráfica infinita simple $G_P = (V, E)$ y así tener un Markov shift. Aunque es posible que existan vértices $v \in V$ tales que E_v o E^v son de cardinalidad infinita (en cuyo caso el σ -espacio pierde la propiedad de ser *localmente compacto*), *vamos a asumir que para todo vértice $v \in V$ y para todo entero $n \geq 1$, hay un número finito de ciclos en v , es decir,*

$$p_n(u) < \infty \quad \Rightarrow \quad q_n(u) < \infty$$

(ver ecuación 1.24). Definimos la *entropía topológica* como $h(X_P) = \log \lambda$, donde

$$\lambda = \limsup_{n \geq 1} |\log p_n(u)|^{1/n} \tag{2.22}$$

(el valor de λ no depende del vértice $u \in V$).

De nuevo, para matrices irreducibles, las definiciones 1, 2 y 3 no dependen del vértice que se escoja (ver sección §2.1.2).

Convención. [ENTROPÍA TOPOLÓGICA FINITA] *Vamos a asumir a lo largo de este trabajo que todos los Markov shift's son de entropía topológica finita.*

Recordemos la siguiente definiciones clásicas en teoría de probabilidad, aunque primero la enunciamos en el contexto de Markov shift's *a la* [3]).

Definición 2.23. Sea A una matriz de adjacencia de un Markov shift X_A irreducible de entropía finita $\log \lambda$. Decimos que X_A es

1. recurrente si $\sum q_n(u)/\lambda^n = 1$,
2. recurrente positiva si es recurrente y $\sum qn_n(u)/\lambda^n < \infty$,
3. recurrente nula si es recurrente pero no recurrente positiva y
4. fuertemente recurrentes positivas si $\limsup |q_n(u)|^{1/n} < \lambda$.

Convención. [CADENAS FUERTEMENTE RECURRENTES POSITIVAS] En este trabajo todos los Markov shift's son fuertemente recurrentes positivos (incluye a todos los SFT's).

La clase de las cadenas de Markov fuertemente recurrentes positivas es la que más asemeja a la clase de los SFT's (entre las caracterizaciones de ser fuertemente positivo recurrente destacan los decaimientos exponenciales, ver [9]). Ahora hay que extender el concepto de medida al contexto numerable.

Definición 2.24. Sea P una matriz estocástica irreducible que representa a una cadena de Markov $(X_n)_{n \in \mathbb{Z}}$. Sea $G_P = (V, E)$ la gráfica con matriz de adjacencia $P^\#$. Decimos que P es recurrente si existe (equivalentemente, para todo) $v \in V$,

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n = v \mid X_0 = v, X_1 \neq v, \dots, X_{n-1} \neq v, X_n = v) = 1$$

y es positiva recurrente si es recurrente y además

$$\sum_{n=1}^{\infty} n \mathbb{P}(X_n = v \mid X_0 = v, X_1 \neq v, \dots, X_{n-1} \neq v, X_n = v) < \infty.$$

Es importante distinguir aquí los dos conceptos que hemos presentado dentro de la clasificación de estados. La Definición 2.23 no involucra a una cadena de Markov en una gráfica, considera únicamente Markov shift's, las probabilidades de transición *no* están involucradas en las definiciones, únicamente el valor de la entropía y el número de ciclos de primer retorno. En contraste, la Definición 2.24 involucra a las probabilidades de transición. La siguiente definición puede ayudar a entender mejor la diferencia.

Definición 2.25. Sea P una matriz estocástica (posiblemente infinita) y sea G la matriz de adjacencia de $P^\#$. Decimos que la cadena de Markov que presenta P tiene esqueleto fuertemente recurrente positivo si X_G es fuertemente recurrente positivo.

Obsérvese que no se involucró en las definiciones al vector de probabilidades iniciales. Abordaremos este tema ahora en la siguiente sección.

2.1.4. Teoría de Perron-Frobenius y distribuciones estacionarias

Como vimos anteriormente, tanto para cadenas de Markov en SFT's por vértices o por aristas, la medida σ -invariante está definida en términos de un vector de probabilidades iniciales. En el caso de SFT's irreducibles, es irrelevante el vector de probabilidades iniciales ya que, a largo plazo, el proceso tiende a ser *estacionario* en el sentido de que la distribución de frecuencias de “visitas” a cada uno de los vértices converge a una *distribución estacionaria*. Este será el caso también para matrices estocásticas que sean irreducibles, no necesariamente finitas, pero sí (fuertemente) recurrentes positivas. Comencemos con la definición de distribución estacionaria.

Definición 2.26. *Sea P una matriz estocástica. Decimos que un vector de probabilidades iniciales π es una distribución estacionaria si*

$$\pi P = \pi.$$

La existencia y unicidad de distribuciones estacionarias de matrices estocásticas irreducibles la determina el Teorema de Perron-Frobenius (Teorema 1.31). A continuación lo enunciamos en el caso particular de matrices estocásticas.

Teorema 2.27 (Perron-Frobenius). *Sea P una matriz estocástica irreducible de tamaño $n \times n$. Entonces*

1. *1 es valor característico de P .*
2. *Si $\lambda \neq 1$ es valor característico de P , entonces $|\lambda| < 1$.*
3. *Existe un único vector característico $\pi = (\pi_1, \dots, \pi_n)$ correspondiente a 1 tal que $\pi_i > 0$ para toda $i = 1, \dots, n$ y $\sum_{i=1}^n \pi_i = 1$.*

Observación 2.28. *Sea P una matriz estocástica irreducible. El vector característico π dado por 3 en el Teorema 2.27 resulta ser la distribución estacionaria de P .*

Entonces si tenemos una matriz de probabilidades de transición P , para conocer la probabilidad estacionaria lo que necesitamos hacer es encontrar el vector característico izquierdo con valor característico 1. Esto es relativamente fácil pues como P es una matriz estocástica, por el Teorema de Perron-Frobenius sabemos que tiene un valor característico de norma máxima, como ya enunciamos, dicho valor es 1. Por lo tanto siempre vamos a tener dicho vector, y para poder determinar cuándo dicho vector es único, tenemos que pedir ciertas condiciones a la matriz P , en este caso, que la cadena de Markov sea irreducible.

La razón de por qué se denomina a π en el Teorema 2.27 distribución *estacionaria* está dada en el siguiente resultado clásico, el cual enunciamos, por simplicidad, para matrices mezcla¹.

¹Una matriz P mezcla si el SFT que determina tiene periodo 1.

Teorema 2.29 (Ver [4, 8]). *Sea P una matriz estocástica irreducible y además que mezcla. Sea π su distribución estacionaria. Entonces*

$$\lim_{n \rightarrow \infty} P_{ij}^n = \pi_j \quad \text{para toda } j \in V.$$

Convención. Siempre que P sea una matriz irreducible, asumiremos a su (única) distribución estacionaria como el vector de probabilidades iniciales.

2.1.5. Entropía de teoría de la medida.

Como vimos en la sección §1.1.4 se puede ver qué tan complicado es un σ -espacio a partir de su entropía topológica. Cuando a dicho espacio le asignamos una probabilidad, entonces la probabilidad de que ciertos bloques sucedan con respecto a otros debe de modificar la entropía del sistema. De manera heurística, es natural pensar que si las probabilidades de transición cambian, entonces habrá bloques o *eventos* que se vuelvan más o menos probables, y esto debería de afectar la entropía. Por lo tanto, sería deseable que el concepto de entropía fuera más sensible a las probabilidades de transición. A continuación damos una definición de *entropía de teoría de la medida* o *entropía de Shannon* de un σ -espacio para cadenas de Markov homogéneas e irreducibles.

Para definir esta nueva entropía de teoría de la medida, primero necesitamos definir algunos conceptos. Primero vamos a definir la entropía de un vector estocástico.

Definición 2.30. *La entropía de un vector estocástico $p = (p_1, \dots, p_k)$ es,*

$$h(p) = - \sum_{i=1}^k p_i \log(p_i)$$

con $0 \log 0 = 0$ por convención.

Haciendo un cálculo simple, para un vector de probabilidad $p = (p_1, \dots, p_k)$, tenemos que

$$0 \leq h(p) \leq \log k \tag{2.31}$$

con

$$h(p) = \log k \text{ si y sólo si } p = (1/k, 1/k, \dots, 1/k)$$

y

$$h(p) = 0 \text{ si y sólo si } p_i = 1 \text{ para alguna } i = 1, \dots, k.$$

Es decir que la entropía de un vector estocástico alcanza su valor máximo cuando las probabilidades de cada una de las entradas son iguales entre si, y su mínimo cuando la probabilidad está concentrada en un estado (punto masa).

Definición 2.32. Sea P una matriz estocástica irreducible que presenta a una cadena de Markov recurrente positiva. Para cada n ponemos las probabilidades de todos los caminos de longitud n en un vector estocástico $p^{(n)}$. La entropía de la teoría de la medida se la cadena de Markov se define como

$$h(X_\mu) = \lim_{n \rightarrow \infty} \frac{1}{n} h(p_A^{(n)}),$$

donde $\mu = \mu_P$ es la medida σ -invariante definida por P .

Es conocido que el límite existe (ver [20]). En teoría de la información, esta noción de entropía es llamada razón de entropía y se atribuye a Shannon (ver [18]). Esta nueva definición de entropía refleja la razón a la cual $\exp(h(p^{(n)}))$ crece.

Es posible escribir a la entropía de la teoría de la medida en términos de las probabilidades de transición y la distribución estacionara de la cadena de Markov, a saber,

$$h(X_\mu) = - \sum_{e \in E(G)} \mu(i(e)) \mu(e | i(e)) \log(\mu(e | i(e)))$$

(la prueba de que la fórmula anterior es equivalente a la de la Definición 2.32; la prueba de esta equivalencia se puede encontrar en [20]).

Concluimos esta sección simplemente observando que la entropía topológica acota superiormente a la entropía de teoría de la medida. Tenemos que si G es una gráfica irreducible, con matriz de adyacencia A , $\lambda = \lambda_A$ el valor Perron y w, v los vectores Perron de A tomados de forma que $v \cdot w = 1$. Luego entonces los caminos de longitud n son aproximadamente $(\sum_{I, J \in V} w_J v_I) \lambda^n$. Incluso, el resultado se reduce para cuando tenemos una n lo suficientemente grande, es decir que hay a lo más $2 \left(\sum_{I, J \in V} w_J v_I \right) \lambda^n$ caminos de longitud n . Del hecho anterior y de la ecuación 2.31 tenemos que

$$h(p^{(n)}) \leq \log \left(2 \sum_{I, J \in V} w_J v_I \lambda^n \right).$$

De forma que,

$$h(X_\mu) \leq \log \lambda.$$

Entonces, como se había afirmado, la entropía de teoría de la medida siempre está acotada superiormente por la entropía del sistema sin medida (la entropía topológica).

2.2. Matrices doblemente estocásticas

En esta sección exponemos los resultados más fundamentales acerca de las matrices doblemente estocásticas. Se dan algunas propiedades y diferentes ejemplos de éstas. En particular, se presenta

la demostración del Teorema de Birkhoff-von Neumann. La intención de la sección es la de brindar un breve panorama de las matrices doblemente estocásticas que consideraremos en el capítulo 4.

Como sabemos las matrices estocásticas son aquellas con entradas no negativas tales que sus renglones suman 1. Si añadimos la propiedad de que también las columnas de la matriz sumen 1, entonces tenemos una matriz doblemente estocástica. Pongamos una definición formal de lo dicho.

Definición 2.33. *Sea P una matriz con entradas no negativas. Si cada columna y cada renglón suman 1, entonces P es una matriz doblemente estocástica.*

Las matrices doblemente estocásticas tienen propiedades interesantes que se verán en esta sección, pero antes demos algunos ejemplos simples de matrices doblemente estocásticas

Ejemplo 2.34. Claramente $P = I$ es doblemente estocástica, pues en cada renglón y cada columna sólo tienen un 1, por lo que la suma nos da 1.

Definición 2.35. *Una matriz π es de permutación, si se obtiene intercambiando los renglones y las columnas de la matriz identidad I .*

Ejemplo 2.36. Una matriz de permutación es una matriz doblemente estocástica.

Denotemos al conjunto de matrices doblemente estocásticas de tamaño $n \times n$ por Ω_n . Ahora enlistemos algunas propiedades de este conjunto.

Proposición 2.37. *El conjunto Ω_n visto como un subconjunto de \mathbb{R}^{n^2} , es cerrado, acotado y convexo.*

PRUEBA. Es sencillo ver que Ω_n es acotado. El que Ω_n es cerrado se sigue de la continuidad de la suma. Demostraremos que Ω_n es convexo. Sean $A, B \in \Omega_n$. Para cualquier $0 \leq \lambda \leq 1$, la matriz $\lambda A + (1 - \lambda)B$ tiene todas las entradas no negativas, y para toda i tenemos que,

$$\lambda \sum_j a_{i,j} + (1 - \lambda) \sum_j b_{i,j} = 1.$$

Por lo que la suma de los renglones nos da 1. De forma análoga tenemos que la suma de las columnas nos da 1. De esta forma demostramos que es convexo.

Observemos el conjunto Ω_2

$$\Omega_2 = \left\{ A \mid A = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix} \text{ con } 0 \leq p \leq 1 \right\}.$$

Es fácil ver que podemos describirlo de la siguiente forma:

$$\Omega_n = \left\{ p \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (1-p) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ con } 0 \leq p \leq 1 \right\}.$$

Entonces tenemos que Ω_2 son todas las posibles combinaciones convexas de dos matrices (de permutación), por lo tanto podemos decir que Ω_2 es el conjunto convexo generado por,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Un punto x_0 de un conjunto convexo S se la llama *punto extremo* de S si $S \setminus \{x_0\}$ es también un conjunto convexo.

A continuación demostraremos un lema, el cual nos ayudará a demostrar que las matrices de permutación son precisamente los puntos extremos del conjunto Ω_n .

Lema 2. *Un punto x_0 de un conjunto convexo S no es un punto extremo de S si y sólo si hay un par de puntos $u, w \in S$ tales que $x_0 = \frac{1}{2}u + \frac{1}{2}w$; esto es que x_0 es el punto medio de un segmento no trivial en S .*

PRUEBA. Supongamos que $x_0 \in S$ no es un punto extremo de S . Entonces $S \setminus \{x_0\}$ no es convexo. Por convención el conjunto vacío y el conjunto con un solo punto son conjuntos convexas. Por lo tanto existen dos elementos $v, u \in S \setminus \{x_0\}$ tales que $\lambda u + (1-\lambda)v \notin S \setminus \{x_0\}$ para alguna $\lambda \in (0, 1)$. Por lo que $\lambda u + (1-\lambda)v = x_0$. Sin pérdida de generalidad, podemos tomar a $\frac{1}{2} < \lambda < 1$. Si $\lambda = \frac{\alpha+1}{2}$, entonces $0 < \alpha < 1$ y $x_0 = \frac{1}{2}u + \frac{1}{2}w$ donde $w = \alpha u + (1-\alpha)v$.

El regreso, si $x_0 = \frac{1}{2}u + \frac{1}{2}w$ y $u \neq w$, entonces $u, w \in S \setminus \{x_0\}$, pero $x_0 \notin S \setminus \{x_0\}$, por lo que $S \setminus \{x_0\}$ no es un conjunto convexo. ✦

Tomando el resultado anterior y la caracterización de Ω_2 , podemos intuir que las matrices de permutación son puntos extremos de Ω_n y que éstos son lo únicos puntos extremos del conjunto Ω_n , pero primero enunciemos la proposición.

Proposición 2.38. *Cualquier matriz de permutación de tamaño $n \times n$ es punto extremo de Ω_n .*

PRUEBA. Sea $\Pi = (\pi_{i,j})$ una matriz de permutación en Ω_n . Si Π no es un punto extremo de Ω_n , por el lema 2, $\Pi = \frac{A+B}{2}$ con $A \neq B$ y $A, B \in \Omega_n$. Ahora bien, para toda $i, j \in \{1, \dots, n\}$, se tiene $\pi_{i,j} = 0$ ó 1 . Como $A, B \in \Omega_n$, tenemos que $0 \leq A_{i,j} \leq 1$ y $0 \leq B_{i,j} \leq 1$. Por lo tanto si $\pi_{i,j} = 0$, entonces $A_{i,j} = B_{i,j} = 0$ y si $\pi_{i,j} = 1$, entonces $A_{i,j} = B_{i,j} = 1$. Esto contradice el hecho que $A \neq B$. ✦

2.2.1. Permanentes

Dada una matriz cuadrada $A = A_{i,j}$ de tamaño n , la *diagonal* asociada a una permutación τ de $\{1, 2, \dots, n\}$, es el conjunto $\{A_{1,\tau(1)}, A_{2,\tau(2)}, \dots, A_{n,\tau(n)}\}$. Una *diagonal* se le dice positiva si cada entrada en la diagonal es positiva. El producto $\prod_{i=1}^n A_{i,\tau(i)}$ se le denomina el *producto diagonal* de A asociado a la permutación τ .

Ejemplo 2.39. Consideremos la siguiente matriz de 4×4 ,

$$\begin{pmatrix} 3 & 5 & 2 & 7 \\ 6 & 0 & 1 & 4 \\ 3 & 2 & 5 & 1 \\ 2 & 9 & 1 & 3 \end{pmatrix}.$$

La diagonal correspondiente a la permutación $\tau: \{1, 2, 3, 4\} \rightarrow \{3, 1, 2, 4\}$ consiste en el conjunto $\{2, 6, 2, 3\} = \{A_{1,3}, A_{2,1}, A_{3,2}, A_{4,4}\}$. El producto diagonal asociado a τ es 72.

La suma de todos los posibles productos diagonales de una matriz es llamado el *permanente* de la matriz. Por lo que si $A = A_{i,j}$ es de tamaño $n \times n$ el permanente de la matriz se define como,

$$\text{per}(A) = \sum_{\tau} \prod_{i=1}^n A_{i,\tau(i)}.$$

Donde la suma es tomada sobre todas las $n!$ permutaciones τ de $\{1, 2, \dots, n\}$.

Modificando un poco la definición del permanente, obtenemos la fórmula del determinante. Es decir, en la suma del permanente sumamos o restamos dependiendo del signo de la permutación. Recordemos que toda permutación se puede escribir como producto de transposiciones, por lo que se sumará cuando la permutación sea un número par de transposiciones y se restará cuando la permutación sea un número impar de transposiciones. Entonces la fórmula del determinante de una matriz queda de la siguiente forma:

$$\det A = \sum_{\tau \in S(n)} \text{sig}(\tau) \prod_{i=1}^n A_{i,\tau(i)}$$

Para demostrar el teorema principal de esta sección, necesitamos primero del siguiente teorema.

Teorema 2.40 (Frobenius-König). *El permanente de una matriz no negativa A de tamaño $n \times n$ es cero si y sólo si A tiene un submatriz de ceros de tamaño $r \times s$ con $r + s = n + 1$.*

PRUEBA. Si $A = 0$, el resultado se da. Sea $A = A_{i,j}$ y $A_{i_0,j_0} \neq 0$ para alguna i_0, j_0 . Por la suposición de que el permanente es cero, entonces todo producto diagonal es cero. Denotamos por $A(i, j)$ a

la submatriz que resulta de eliminar el i -ésimo renglón y la j -ésima columna de A . Como todo permanente es cero, si tomamos la submatriz $A(i_0, j_0)$ entonces $\text{per}(A(i_0, j_0)) = 0$. Aplicando el mismo argumento y utilizando inducción sobre la matriz $A(i_0, j_0)$, podemos decir que A tiene una submatriz de orden $u \times v$ tal que $u + v = n$. Podemos encontrar matrices de permutación Π_1 y Π_2 tales que,

$$\Pi_1 A \Pi_2 = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}.$$

Donde B y D son matrices cuadradas de orden u y $n - u$, respectivamente. Como $\text{per}(A) = 0$, entonces $\text{per}(B) = 0$ o $\text{per}(D) = 0$. Supongamos que $\text{per}(B) = 0$. Por inducción B tiene una submatriz de tamaño $p \times q$ con $p + q = u + 1$. Sin pérdida de generalidad, podemos asumir que dicha matriz está formada por los primeros p renglones y las últimas q columnas de B . Entonces A tiene una matriz de ceros de tamaño $p \times (q + n - u)$ y $p + q + n - u = u + 1 + n - u = n + 1$. Un argumento similar se aplica cuando $\text{per}(D) = 0$.

Para probar el regreso, supongamos, sin pérdida de generalidad, que la matriz formada por los primeros r renglones y las primeras s columnas de A es la submatriz de ceros, donde $r + s = n + 1$. Sea τ cualquier permutación del conjunto $\{1, 2, \dots, n\}$. Si mostramos que $A_{i, \tau(i)} = 0$ para alguna i acabamos, pues todo el producto diagonal sería igual a 0 para toda permutación τ . Supongamos que $A_{i, \tau(i)} \neq 0$ para toda i , entonces tenemos que

$$\{\tau(1), \tau(2), \dots, \tau(r)\} \cap \{1, 2, \dots, s\} = \emptyset$$

lo que contradice el hecho que $r + s = n + 1$ por lo tanto $A_{i, \tau(i)} = 0$ para alguna i demostrando así el regreso del teorema. ✦

Entonces ¿cómo se comporta el permanente de una matriz cuando ésta es doblemente estocástica? La respuesta a tal pregunta la tenemos como lema, ya que éste nos ayudará en la prueba del teorema principal de la sección.

Lema 3. *Si A es una matriz doblemente estocástica, entonces A tiene una diagonal positiva.*

PRUEBA. Sea A una matriz doblemente estocástica de tamaño $n \times n$. Si demostramos que $\text{per}(A) > 0$ terminamos nuestra prueba, pues el producto de alguna diagonal sería positiva y esto sólo sucede en una matriz no negativa cuando la diagonal es positiva. Supongamos que $\text{per}(A) = 0$, entonces por el Teorema 2.40, A tiene una submatriz de tamaño $r \times s$ con $r + s = n + 1$. Podemos asumir, que después de permutar los renglones y las columnas de A si es necesario, de forma que,

$$A = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

donde la matriz cero es de tamaño $r \times s$ con $r + s = n + 1$. Entonces B es de tamaño $r \times (n - s)$ y D es de tamaño $(n - r) \times s$. Como A es doblemente estocástica, los renglones de B suman 1 y

entonces las entradas de B suman a lo más r . La suma de las columnas de A también son iguales a 1 y por lo tanto las entradas de B junto con las entradas de C suman a lo más $n - s$. Por lo que tenemos que tener que $r \leq n - s$, la cual es una contradicción por que $r + s = n + 1$. Completando la prueba. ✦

2.2.2. El Teorema de Birkhoff-Von Neumann

El siguiente teorema es el principal de la sección actual, en este lo que demostramos es que: cualquier matriz doblemente estocástica está dentro de un politopo, el cual tiene como puntos extremos a las matrices de permutación y está generado por algunas de éstas.

Teorema 2.41 (Birkhoff-Von Neumann). *Cualquier matriz A doblemente estocástica puede ser escrita como una combinación convexa de un número finito de matrices de permutación; esto es*

$$A = \lambda_1 \Pi_1 + \lambda_2 \Pi_2 + \cdots + \lambda_m \Pi_m.$$

Donde $\Pi_1, \Pi_2, \dots, \Pi_m$ son matrices de permutación y $0 \leq \lambda_1, \lambda_2, \dots, \lambda_m \leq 1$ con $\sum_{i=1}^m \lambda_i = 1$.

PRUEBA. Sea A una matriz doblemente estocástica de tamaño $n \times n$. Si A es una matriz de permutación, entonces la prueba es obvia. De otra forma tenemos que por el Lema 3, A tiene una diagonal positiva. Sea Π_1 la permutación de tal diagonal. Si $\{A_{1,r_1}, A_{1,r_2}, \dots, A_{n,r_n}\}$ es la diagonal positiva (de forma que $r_i = \Pi_1(i)$), entonces

$$(\Pi_1)_{ij} = \begin{cases} 1 & \text{si } (i, j) \in \{(1, r_1), (1, r_2), \dots, (n, r_n)\} \\ 0 & \text{cualquier otro caso.} \end{cases};$$

Sea $\lambda_1 = A_{t,r_t} = \min\{A_{1,r_1}, A_{1,r_2}, \dots, A_{n,r_n}\} > 0$. Es claro que $\lambda_1 < 1$. Tenemos que

$$A = \lambda_1 \Pi_1 + (1 - \lambda_1) \left\{ \frac{1}{1 - \lambda_1} (A - \lambda_1 \Pi_1) \right\}.$$

Es claro que $B = \frac{1}{(1 - \lambda_1)} (A - \lambda_1 \Pi_1)$ es de nuevo una matriz doblemente estocástica y al menos tiene una entrada cero más que A . Si B es una matriz de permutación, entonces $A = \lambda_1 \Pi_1 + (1 - \lambda_1) B$ y terminamos. De otra forma podemos encontrar una matriz de permutación Π_2 , como la anterior, de forma que $B = \mu_2 \Pi_2 + (1 - \mu_2) C$ donde C es una matriz doblemente estocástica. Si C es una matriz de permutación, entonces $A = \lambda_1 \Pi_1 + \lambda_2 \Pi_2 + \lambda_3 \Pi_3$ donde $\lambda_2 = (1 - \lambda_1) \mu_2$, $\lambda_3 = (1 - \lambda_2) (1 - \mu_2)$ y $\Pi_3 = C$. Más aun λ_1, λ_2 y λ_3 son no negativas y su suma es 1. También nótese que C tiene al menos dos ceros más que A . Si C no es una matriz de permutación podemos continuar este proceso como antes con otra iteración. Entonces con a lo más $n^2 - n$ iteraciones, nuestro algoritmo tiene que terminar con una matriz de permutación y por lo tanto A es una combinación convexa de a lo más $n^2 - n + 1$ matrices de permutación.



Entonces hemos demostrado que cualquier matriz doblemente estocástica $A \in \Omega_n$ puede ser expresada como una combinación convexa de a lo más $n^2 - 2n - 2$ matrices de permutación. Este hecho se puede deducir utilizando el Teorema de Carathéodory, el cual dice que si $S \subset \mathbb{R}^n$ es un conjunto convexo, entonces cualquier punto en S puede ser expresado como combinación convexa de a lo más $n + 1$ puntos de S .

Capítulo 3

El invariante del árbol generador

En este capítulo vamos a definir al *invariante del árbol generador*, una de las dos clases de objetos centrales de estudio en esta Tesis. La otra clase de objetos, los sistemas de ciclos de primer retorno, la vamos a posponer hasta el capítulo 4, endonde veremos al invariante del árbol generador en el contexto de los sistemas de ciclos de primer retorno.

3.1. El invariante del árbol generador

Definición 3.1. Sea $G = (V, E)$ una gráfica dirigida. Un árbol en G enraizado en $r \in V$ es una subgráfica T de G , tal que todo vértice en T excepto r tiene una única flecha de salida en T , r no tiene flecha de salida en T y para cada vértice de T distinto a r , existe una única trayectoria dirigida en T que termina en r . Entonces decimos que el árbol T está enraizado en r . Un árbol T es generador si contiene todos los vértices de G .

En la Figura 3.1 vemos una gráfica y un árbol generador de la gráfica con raíz en el vértice 5.

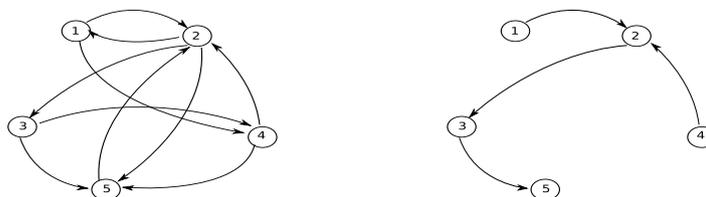


Figura 3.1: Una gráfica y uno de sus árboles generadores enraizado en el vértice 5.

Dada una gráfica dirigida $G = (V, E)$, sea \mathcal{S}_r el conjunto de árboles generadores con raíz r de la gráfica G , y $\mathcal{S} = \bigcup_{r \in V} \mathcal{S}_r$ es el conjunto de todos los árboles generadores en G .

Sea P una matriz estocástica. Sea $G_P = (V; E)$ la gráfica con matriz de adjacencia $P^\#$. Para una arista $e \in E$, $\mathbb{P}(e) = P_{i(e)t(e)}$ es probabilidad de transición de la arista e , Dada una subgráfica H de G , definimos el *peso* de H como

$$w_P(H) = \prod_{e \in H} \mathbb{P}(e). \quad (3.2)$$

Definición 3.3. Sea P una matriz estocástica y \mathcal{S} el conjunto de árboles generadores de $G_P = (V, E)$. Se define el invariante del árbol generador de P como

$$\tau(P) = \sum_{S \in \mathcal{S}} w_P(S). \quad (3.4)$$

Definimos también, para cada $v \in V$, el invariante del árbol generador local en v como

$$\tau_v(P) = \sum_{S \in \mathcal{S}_v} w_P(S). \quad (3.5)$$

Es claro de la definición que

$$\tau(P) = \sum_{v \in V} \tau_v(P).$$

En caso de que la matriz P sea infinita, las ecuaciones (3.2), (3.4) y (3.5) hacen sentido *únicamente* como límites.

Ejemplo 3.6. Tomemos la siguiente matriz

$$P = \begin{pmatrix} 0 & 1/3 & 2/3 \\ 2/3 & 0 & 1/3 \\ 1/3 & 2/3 & 0 \end{pmatrix}.$$

Tenemos que hay 3 árboles con raíz vértice 1 lo que se muestra a continuación.

En la Figura 3.2 se muestra la gráfica determinada por la matriz P , entonces calculemos el peso de los árboles enraizados en el vértice 1. El peso del primero es $\frac{2}{9}$, el peso del segundo es $\frac{4}{9}$ y el tercero es de $\frac{1}{9}$. Hacemos lo mismo para cada vértice y sumamos, por lo que el invariante del árbol generador de la gráfica es $\frac{21}{9}$.

•

La siguiente y última sección de este capítulo está dedicada a la demostración de que el invariante del árbol generador es un invariante bajo conjugación.

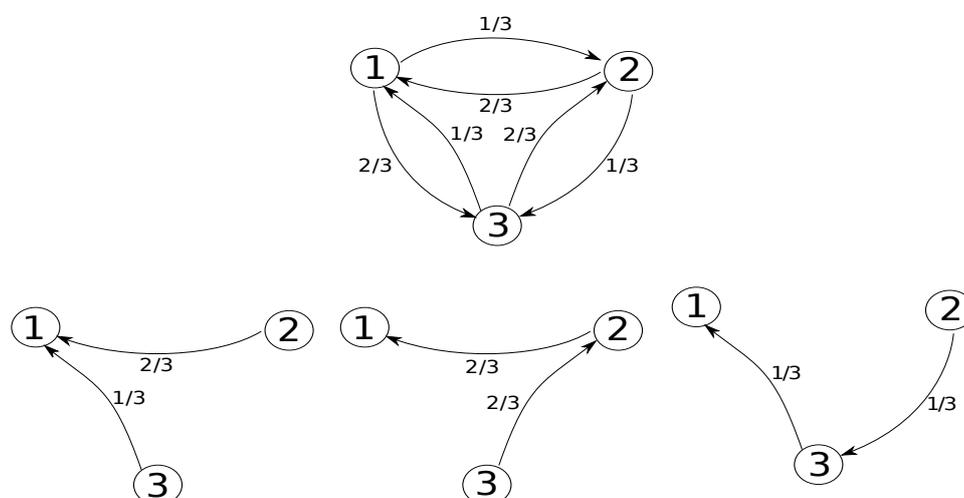


Figura 3.2: La gráfica determinada por la matriz P y los 3 árboles generadores en el vértice 1.

3.2. Invarianza bajo conjugación del invariante de árbol generador

En esta sección presentamos la demostración del hecho de que el invariante del árbol generador es invariante bajo conjugaciones, es decir, bajo isomorfismo de bloques. La demostración que presentamos se basa en la prueba que Lind y Tuncel hicieron en [13].

Teorema 3.7. *Sean P y Q matrices estocásticas irreducibles tales que los σ -espacios de las matrices son isomorfos bajo isomorfismos de bloques (conjugados), entonces $\tau(P) = \tau(Q)$.*

PRUEBA. Primero hagamos la prueba para la operación de in-splitting sobre un vértice k y consideremos la partición en sólo 2 conjuntos, es decir $E^k = \mathcal{F}_1 \cup \mathcal{F}_2$.

Sean G y G' las gráficas original y la que resulta de hacer un in-splitting en el vértice k . Para la prueba necesitamos una relación entre cierto conjunto de árboles generadores en G y un conjunto similar de árboles generadores en G' . Para tal propósito consideremos la terna de la forma (T, T_1, T_2) donde T es un árbol en G enraizado en un vértice r y cada T_i es un árbol en G enraizado en k usando las aristas de \mathcal{F}_i ($i = 1, 2$). Los tres árboles son ajenos salvo en el vértice k y contienen a todos los vértices de G . Se da la posibilidad de que T sea el árbol vacío o bien el árbol con un solo vértice.

Cada una de estas ternas (T, T_1, T_2) en G le corresponde una terna (T', T'_1, T'_2) en G' donde T es copiado exactamente en T' y T'_m es el árbol enraizado en k_m que viene de T_m ($m = 1, 2$). Tal terna tiene la propiedad que T' es un árbol enraizado en r , T'_m es un árbol enraizado en k_m para $m = 1, 2$, los tres árboles son ajenos, y si el árbol original en G contenía a todos los vértices,

entonces estos tres árboles contienen a todos los vértices de G' . Esta relación entre las ternas de G y las ternas de G' es claramente uno a uno.

Sea $\mathcal{S}(T, T_1, T_2)$ el conjunto de todos los árboles generadores en G enraizados en r que contengan a T , T_1 y T_2 . De forma similar definimos a $\mathcal{S}(T', T'_1, T'_2)$ en G' . Es claro que el conjunto $\mathcal{S}(G)$ de árboles generadores es la unión ajena de $\mathcal{S}(T, T_1, T_2)$ sobre todas las ternas posibles, de forma similar $\mathcal{S}(G')$ es la unión ajena de $\mathcal{S}(T', T'_1, T'_2)$.

Entonces para probar que $\tau(P) = \tau(P')$ es suficiente probar que

$$\sum_{S \in \mathcal{S}(T, T_1, T_2)} w_P(S) = \sum_{S' \in \mathcal{S}(T', T'_1, T'_2)} w_{P'}(S').$$

Fijemos una terna (T, T_1, T_2) . Para tener un árbol generador en G enraizado en r que contenga estos tres árboles tenemos que añadir una arista de k a cualquiera de los vértices del árbol T . Si denotamos por $p(k, T)$ a la suma de todas las posibles probabilidades de transición de k a los vértices de T . Entonces tenemos que

$$\sum_{S \in \mathcal{S}(T, T_1, T_2)} w_P(S) = w_P(T)w_P(T_1)w_P(T_2)p(k, T). \quad (3.8)$$

La igualdad se da porque por cada arista que tomemos de k hacia un vértice de T es un árbol generador distinto y se sigue del hecho que el peso del árbol es el producto de sus aristas.

Ahora consideremos la terna correspondiente en G' , (T', T'_1, T'_2) . En G' tenemos tres formas de tener un árbol generador con raíz r . Unir k_1 con el árbol T'_2 y k_2 con el árbol T' , o unir k_2 con el árbol T'_1 y a k_1 con el árbol T' , o bien unir k_1 y k_2 con el árbol T' .

Veamos cómo afecta esta forma de unir a los tres árboles al peso del árbol generador en G' . En cada caso sólo estamos agregando dos aristas a nuestros árboles, por lo que estamos agregando $P'(k_1, T'_2)P'(k_2, T')$ al peso de los árboles en el primer caso. De igual forma $P'(k_2, T'_1)P'(k_1, T')$ y $P'(k_1, T')P'(k_2, T')$ para el segundo y tercer caso respectivamente.

Por lo que la suma de los pesos de los árboles generadores queda de la siguiente forma

$$\sum_{S' \in \mathcal{S}'(T', T'_1, T'_2)} w_{P'}(S') = w_{P'}(T')w_{P'}(T'_1)w_{P'}(T'_2) \times [P'(k_1, T'_2)P'(k_2, T') + P'(k_2, T'_1)P'(k_1, T') + P'(k_1, T')P'(k_2, T')]. \quad (3.9)$$

Para ver que las ecuaciones (3.8) y (3.9) son iguales tenemos que ver como queda la matriz P' con respecto a la matriz P .

Para definir a P' vamos a definirla elemento a elemento. Como la partición es en dos conjuntos la matriz tiene $|V(G)| + 1$ columnas y $|V(G)| + 1$ renglones. Denotemos $P'(i, j)$ al elemento en el

renglón i y en la columna j de la matriz P' . Ahora bien, si $i \neq k_m$ y $j \neq k_m$ con $m = 1, 2$, entonces $P'(i, j) = P(i, j)$ pues la gráfica G fuera de los vértices que son adyacentes a k se mantiene igual. Si i no tiene flecha hacia k en G , entonces $P'(i, k_m) = 0$ con $m = 1, 2$, pues sigue sin tener adyacencia con k_1 o k_2 en G' . Si i tiene flecha e hacia k entonces su peso se va hacia su respectivo vértice, es decir, e pertenece a algún \mathcal{F}_m con $m = 1, 2$. Supongamos que $e \in \mathcal{F}_1$, entonces el peso va sumado en el término $P'(i, k_1)$. Si $i = k_m$ con $m = 1, 2$ y $j \neq k_m$ con $m = 1, 2$, entonces $P'(i, j) = P'(k_m, j) = P(k, j)$, pues las flechas de salida de los vértices k_m con $m = 1, 2$, son las mismas que las del vértice k . Si $i = k_m$ con $m = 1, 2$ y $j = k_m$ con $m = 1, 2$ si k no tiene un lazo entonces $P'(i, j) = 0$, si hay un lazo e , tal e esta en algún \mathcal{F}_m con $m = 1, 2$, sin pérdida de generalidad supongamos que $e \in \mathcal{F}_1$ entonces hay un lazo en k_1 y una arista de k_2 a k_1 , el peso de la arista e es sumado en las respectivas entradas.

Si tenemos un lazo en k , sin pérdida de generalidad dicho lazo estará en \mathcal{F}_1 . Ahora veamos cuál es el peso de los diferentes tipos de árboles en las dos gráficas, $w_P(T) = w_{P'}(T')$, $w_P(T_1) = w_{P'}(T'_1)$ y $w_P(T_2) = w_{P'}(T'_2)$. Dichas igualdades salen directamente de la definición del peso de los árboles y de cómo queda P' con respecto a P . También se tiene las siguientes igualdades $P'(k_1, T') = P'(k_2, T') = P(k, T)$ y $P'(k_1, T'_2) = P(k, T_2) - P(k, k)$, $P'(k_2, T_1) = P(k, T_1)$. Por lo que

$$\begin{aligned} P'(k_1, T'_2)P'(k_2, T') + P'(k_2, T'_1)P'(k_1, T') + P'(k_1, T')P'(k_2, T') = \\ = P(k, T)[P(k, T_1) + P(k, T_2) - P(k, k) + P(k, T)] = P(k, T). \end{aligned}$$

Como T_1 y T_2 son ajenos excepto en k se demuestra la igualdad entre (3.8) y (3.9). Por lo tanto es invariante bajo los in-splitting.

Ahora demostremos que τ es invariante bajo el out-splitting. Para tal motivo veamos cómo queda la matriz con respecto a un out-splitting. Sea G una digráfica con matriz de transición P estocástica. Vamos hacer el out-splitting en un vértice k , tenemos E_k es el conjunto de las aristas que salen del vértice k . Supongamos que la partición de nuestros vértices es en dos conjuntos, es decir $E_k = \mathcal{F}_1 \cup \mathcal{F}_2$ y $\mathcal{F}_1 \cap \mathcal{F}_2 = \emptyset$. Para facilidad de la prueba pensemos que si k tiene un lazo, éste se encuentra en \mathcal{F}_1 .

Sea G' la digráfica que resulta de hacer la operación de out-splitting en el vértice k en G . La matriz estocástica resultante la denotaremos por P' . Ésta está definida como sigue: para los vértices $i \neq k_m$ y $j \neq k_m$ con $m = 1, 2$ tenemos que $P'(i, j) = P(i, j)$ (es fácil de ver tal afirmación, pues las adyacencias de dichos vértices se mantienen igual).

Para definir los casos restantes definimos a $q = \sum_{e \in \mathcal{F}_1} p(e)$, por lo que $1 - q = \sum_{e \in \mathcal{F}_2} p(e)$. Si $i \neq k$ y $j = k_m$ con $m = 1$, tenemos que $P'(i, j) = P'(i, k_1) = qP(i, k)$ con $m = 2$ $P'(i, j) = P'(i, k_2) = (1 - q)(P(i, k))$ ésto se hace para que en la nueva gráfica se mantenga como una cadena de Markov. De igual forma si $j \neq k_m$ y $i = k_1$ con $m = 1$ tenemos que $P'(i, j) = P'(k_1, j) = P(k, j)/q$. Si $m = 2$ $P'(i, j) = P'(k_2, j) = P(k, j)/(1 - q)$. Para el caso de un lazo tenemos que $P'(k_1, k_1) = P(k, k)$ y $P'(k_1, k_2) = (1 - p)P(k, k)/q$.

Ya que sabemos como queda P' con respecto a los términos de P continuemos con la demostración. Tomemos una pareja (T, U) donde T, U son subgráficas de G , T un árbol enraizado en un vértice r , U es un árbol en G enraizado en k . T y U son ajenos y contienen a todos los vértices de G . Para cualquiera de estos pares denotemos por \mathcal{B} al conjunto de vértices inmediatos a k en U , es decir los vértices que tienen una flecha hacia k y tanto la flecha como el vértice están en U . Cada conjunto $B \subset \mathcal{B}$ induce dos subárboles $U_1(B)$ y $U_2(B)$ donde ambos están enraizados en k y juntos generan a U . $U_1(B)$ es un subárbol de U donde todos los vértices que están en el conjunto B estarán en el árbol $U_1(B)$, mientras que U_2 es el árbol que contiene a los vértices del conjunto $B^c = \mathcal{B} \setminus B$.

Para cada $B \subset \mathcal{B}$ tenemos una tripleta $(T', U'_1(B), U'_2(B))$ en G' , donde T' es la copia de T , $U'_1(B)$ es un árbol enraizado en k_1 usando las flechas de $U_1(B)$ y $U'_2(B)$ es un árbol enraizado en k_2 usando las aristas de $U_2(B)$. Cada par (T, U) corresponde a una colección de tripletas $\{(T', U'_1(B), U'_2(B)) \mid B \subset \mathcal{B}\}$.

Sea $\mathcal{S}(T, U)$ el conjunto de árboles generadores en G enraizados en r que contienen a los subárboles T y U . $\mathcal{S}(T', U'_1(B), U'_2(B))$ el conjunto de árboles generadores en G' que contienen a T' , $U'_1(B)$ y a $U'_2(B)$ como subárboles. Entonces $\mathcal{S}(G)$ es la unión ajena de $\mathcal{S}(T, U)$ y $\mathcal{S}(G')$ es la unión ajena de $\mathcal{S}(T', U'_1(B), U'_2(B))$. Por lo tanto es suficiente probar que para cada par (T, U) tenemos lo siguiente:

$$\sum_{S \in \mathcal{S}(T, U)} w_P(S) = \sum_{B \subset \mathcal{B}} \sum_{S \in \mathcal{S}(T', U'_1(B), U'_2(B))} w_{P'}(S').$$

Fijemos una pareja (U, T) . Sea $a_1 = P(\mathcal{F}_1, T) = \sum \{p(e) \mid e \in \mathcal{F}_1 \text{ y } e \text{ termina en } T\}$, de forma similar definimos a $b_1 = P(\mathcal{F}_1, U)$, $a_2 = P(\mathcal{F}_2, T)$ y $b_2 = P(\mathcal{F}_2, U)$. Tenemos las siguientes igualdades $a_1 + b_1 = q$ y $a_2 + b_2 = 1 - q$.

Para formar un árbol generador en G a partir de la pareja (T, U) necesitamos una arista de k a T , por lo que el valor de la suma de los pesos de los posibles árboles generadores S con respecto a la pareja (T, U) queda como

$$\sum_{S \in \mathcal{S}(T, U)} w_P(S) = w_P(T)w_P(U)P(k, T) = w_p(T)w_p(U)[a_1 + a_2].$$

Tomemos a $B \subset \mathcal{B}$, y formamos la terna $(T', U_1(B), U_2(B))$. Para formar un árbol generador usando la terna tenemos tres formas de hacerlo. Unir al vértice k_1 con $U'_1(B)$ y a k_2 con el subárbol T' , unir k_2 al subárbol $U'_1(B)$ y al vértice k_1 unirlo al subárbol T' , o bien unir a los vértice k_1 y k_2 al subárbol T' .

Entonces la suma de los pesos de los árboles generadores que contienen a esta terna como

subárboles queda de la siguiente forma

$$\sum_{S' \in \mathcal{F}(T', U_1(B), U_2'(B))} w_{P'}(S') = w_{P'}(T') w_{P'}(U_1'(B)) w_{P'}(U_2'(B)) \Phi(B)$$

donde $\Phi(B) = P'(\mathcal{F}_1, U_2'(B))P'(k_2, T') + P'(\mathcal{F}_2, U_1(B))P'(k_1, T') + P'(k_1, T')P'(k_2, T')$.

Es claro que $w_P(T) = w_{P'}(T')$. Sea $n = |B|$, como $U_1(B)$ usa $|B|$ flechas que llegan pues $U_1'(B)$ es un árbol entonces tenemos que cada una de las aristas ha sido multiplicado por q , en el caso de $U_2'(B)$ usa $n - |B|$ aristas y han sido multiplicadas por $1 - q$ entonces tenemos que

$$w_{P'}(U_1'(B)) w_{P'}(U_2'(B)) = q^{|B|} (1 - q)^{n - |B|} w_P(U).$$

Con estas igualdes sólo nos falta demostrar que

$$a_1 + a_2 = \sum_{B \subset \mathcal{B}} q^{|B|} (1 - q)^{n - |B|} \Phi(B).$$

Ahora $P'(k_1, T') = a_1/q$ y $P'(k_2, T') = a_2/(1 - q)$. Denotemos por $\mathcal{F}_1(U)$ al conjunto de aristas en \mathcal{F}_1 que terminan en U . Intercambiando el orden de la suma podemos ver que

$$\begin{aligned} \sum_{B \subset \mathcal{B}} q^{|B|} (1 - q)^{n - |B|} P'(\mathcal{F}_1, U_2'(B)) &= \sum_{e \in \mathcal{F}_1(U)} P'(e) \sum_{e \in B^c} q^{|B|} (1 - q)^{n - |B|} \\ &= \sum_{e \in \mathcal{F}_1(U)} \frac{P(e)}{q} \sum_{k=0}^{n-1} \binom{n-1}{k} q^k (1 - q)^{n-k} = \frac{1-q}{q} \sum_{e \in \mathcal{F}_1(U)} P(e) = \frac{1-q}{q} b_1 \end{aligned}$$

De forma análoga tenemos que

$$\sum_{B \subset \mathcal{B}} q^{|B|} (1 - q)^{n - |B|} P'(\mathcal{F}, U_1'(B)) = \frac{q}{1 - q}.$$

Como $b_1 = 1 - a_1$ y $b_2 = 1 - a_2$ obtenemos que

$$\begin{aligned} \sum_{B \subset \mathcal{B}} q^{|B|} (1 - q)^{n - |B|} \Phi(B) &= \frac{1 - q}{q} b_1 \frac{a_2}{1 - q} + \frac{q}{1 - q} b_2 \frac{a_1}{q} + \frac{a_1}{q} \frac{a_2}{1 - q} \\ &= \frac{a_2(q - a_1)}{q} + \frac{a_1(1 - q - a_2)}{1 - q} + \frac{a_1 a_2}{q(1 - q)} = a_1 + a_2. \end{aligned}$$

Hasta ahora hemos probado que la suma de los pesos de los árboles generadores es invariante bajo las operaciones de in-splitting y out-splitting. Para demostrar que es invariante bajo conjugación, faltaría demostrar que si dos cadenas de Markov son conjugadas entonces podemos obtener una de las gráficas que determinan a un SFT a partir de la otra gráfica, a través de operaciones de splittings. El resultado es entonces una consecuencia del del Corolario 1.46 del Teorema de Descomposición, el cual se aplica después de considerar la presentación en bloques.



(Hay que notar que el Corolario 1.46 así como el Teorema de Descomposición están únicamente enunciados para SFT's y no para cadenas de Markov en SFT's; las extensiones al caso general de cadenas de Markov han sido realizadas por B. Parry y S. Tuncel particularmente, aquí sólo se consideró el caso de los SFT's para así exhibir de manera más sencilla las ideas fundamentales de la demostración del Teorema de Descomposición.)

Ahora nuestro objetivo es ver qué pasa en el caso en que los splittings no sean un número finito, es decir ¿se seguirá cumpliendo que el invariante del árbol generador sea el mismo, bajo un número infinito de splittings? Como veremos en el siguiente capítulo, el invariante del árbol generador no es invariante cuando tenemos un número infinito de splittings.

Capítulo 4

Sistemas de ciclos de primer retorno

Este capítulo es el más importante de esta Tesis porque aquí es que se desarrollan los resultados que obtuvimos y que describimos ya en el resumen de la Tesis en un principio. Comencemos entonces, recordando los sistemas de ciclos de primer retorno.

4.1. Sistemas de ciclos de primer retorno

En esta sección, presentamos a la definición formal de los *sistemas de ciclos de primer retorno*, que son la única clase de Markov shift's que vamos a considerar en este trabajo que no son SFT's.

Sea P una matrix estocástica, finita e irreducible, la cual determina una gráfica $G_P = (V, E)$ con matriz de incidencia $P^\#$. Para cada entero $n \geq 1$ y cada número real $\rho \in [0, 1]$, sean

$$p_n^{(\rho)}(u) = |\{\gamma \mid \gamma \text{ es un ciclo en } u \text{ tal que } \mathbb{P}(\gamma) = \rho\}|$$

y

$$q_n^{(\rho)}(u) = |\{\gamma \mid \gamma \text{ es un ciclo de primer retorno en } u \text{ tal que } \mathbb{P}(\gamma) = \rho\}|$$

de forma que

$$p_n(u) = \sum_{\rho \in [0,1]} p_n^{(\rho)}(u)$$

y

$$q_n(u) = \sum_{\rho \in [0,1]} q_n^{(\rho)}(u).$$

El *sistema de ciclos de primer retorno a v* es una cadena de Markov en los vértices de una gráfica $H = (V_H, E_H)$ que tiene un vértice distinguido que denotamos por \bar{v} y para cada $\rho \in [0, 1]$ y

$n \geq 1$, $q_n^{(\rho)}(v)$ ciclos de longitud n en \bar{v} , ajenos en vértices excepto por el vértice distinguido \bar{v} . Las probabilidades transición son como sigue. Por cada uno de los $q_n^{(\rho)}(v)$ ciclos γ de primer retorno a v en G de longitud n y probabilidad $\mathbb{P}(\gamma) = \rho$, consideremos su ciclo $\bar{\gamma}$ correspondiente en H . A la única arista en $\bar{\gamma}$ que sale de \bar{v} le asociamos la probabilidad de transición $\mathbb{P}(\gamma)$, y al resto de las aristas de $\bar{\gamma}$ les asociamos la probabilidad de transición 1.

Podemos definir a un sistemas de ciclos de primer retorno a través de una serie de potencias

$$f(t) = a_1 t + a_2 t^2 + a_3 t^3 \cdots = \sum_{n=1}^{\infty} a_n t^n \quad (4.1)$$

donde cada $a_n = \rho_1^n + \rho_2^n + \cdots + \rho_k^n$ y ρ_j^n es el peso de la arista que sale de r en H , para el j -ésimo camino cerrado de longitud n . De forma que podemos escribir

$$f(t) = \sum_{n=1}^{\infty} \sum_{\rho \in [0,1]} q_n^{(\rho)}[\rho] t^n \quad (4.2)$$

donde

$$a_n = \sum_{\rho \in [0,1]} q_n^{(\rho)}[\rho].$$

Aquí la serie de potencia es formal (ver Ejemplo 1.15).

Una forma de encontrar dicha serie de potencias a través de nuestra gráfica original G es utilizado la matriz de probabilidades que determina al sistema original. Tomando dicha matriz P podemos determinar el sistema de ciclos de primer retorno usando la siguiente formula:

$$1 - f(t) = \frac{\det(I - Pt)}{\det(I - Q^{(v)}t)}$$

donde $Q^{(v)}$ es la matriz resultante de remover el renglón y la columna que representan¹ al vértice v . Esta fórmula se da utilizando las Proposiciones (1.22) y (1.26).

En la Figura 4.1 tiene dos ciclos de longitud 1, tres de longitud 2, cuatro ciclos de longitud 3, etc.

Otra forma de obtener el sistemas de ciclos de primer retorno, es a través de una sucesión infinita de splittings. Para ver ésto demos un ejemplo sencillo.

Ejemplo 4.3. Consideremos nuestro σ -espacio completo en dos símbolos o “2-shift” (por ahora no consideramos cadenas de Markov). Veamos cuál es la serie de potencias del sistema de ciclos de

¹Aquí es importante interpretar a las entradas de las matrices como series de potencias con coeficientes en el grupo integral sobre el grupo multiplicativo de los números reales.

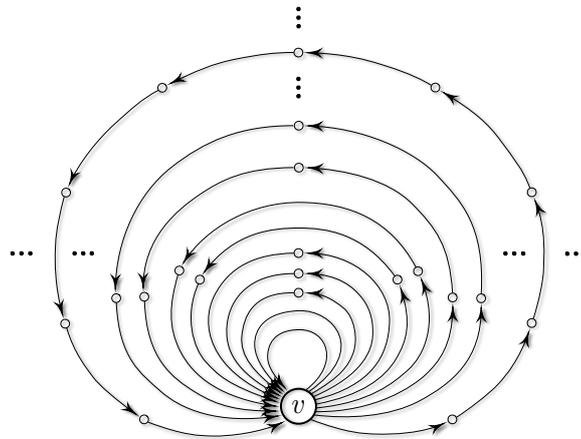


Figura 4.1: Un sistema de ciclos o ‘Loop system’.

primer retorno para nuestro σ -espacio en el v3rtice 1. Tenemos que

$$1 - f(t) = \frac{\det(I - At)}{\det(I - Q^{(1)}t)} = \frac{\left| \begin{pmatrix} 1-t & -t \\ -t & 1-t \end{pmatrix} \right|}{|1-t|} = \frac{(1-t)^2 - t^2}{1-t}$$

de forma que

$$1 - f(t) = \frac{1 - 2t + t^2 - t^2}{1-t} = \frac{1 - 2t}{1-t} = \frac{1-t}{1-t} - \frac{1}{1-t} = 1 - \frac{1}{1-t}.$$

Utilizando la serie geom3trica obtenemos la serie deseada

$$f(t) = t + t^2 + t^3 + t^4 + \dots \tag{4.4}$$

Veamos que la serie cumple la afirmaci3n, es decir que el coeficiente del t3rmino i corresponde al n3mero de caminos de longitud i que tienen un primer retorno al v3rtice 1. Esto es f3cil de verificar utilizando la gr3fica en la Figura 4.2.

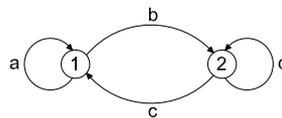


Figura 4.2: El ‘2-shift’

Mostremos como podemos llegar de la gráfica en la Figura 4.2 al sistema de ciclos de primer retorno descrito por la serie (4.4). La forma de hacerlo es a través de una sucesión de splittings.

Al vértice 1 en Figura 4.2 no se le aplica ningún tipo de operación. Aplicamos la operación de in-splitting al vértice 2, con particiones $\{b\}$ y $\{d\}$. La gráfica resultante la tenemos en la parte izquierda de la Figura 4.3. El siguiente paso será hacer un out-splitting al vértice que recibe la arista $\{b\}$, con las particiones $\{c\}$ y $\{d\}$. Lo que nos da es la gráfica del lado derecho de la Figura 4.3. Cabe observar que ya tenemos el ciclo de longitud 2, como se vera a continuación este ya no se toca. De los otros vértices saldrán los ciclos de las distintas longitudes.

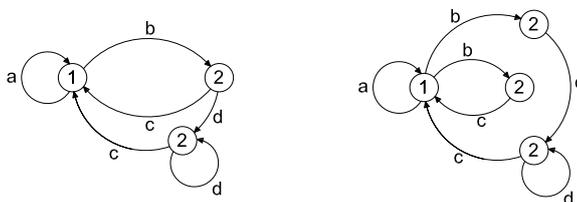


Figura 4.3: Paso 1,2 para generar el sistema de ciclos de primer retorno

Ahora bien al vértice que es vértice final de las aristas $\{d, d\}$ y vértice inicial de la arista $\{c\}$ y al lazo $\{d\}$, le aplicamos la operación de in-splitting con las particiones $\{d\}$ y $\{d\}$. Como resultado obtenemos la gráfica de la izquierda en la Figura 4.4. A dicha gráfica resultante se le aplica un out-splitting en el vértice que es vértice final solo de la arista $\{d\}$, vértice inicial de las aristas $\{c, d\}$ y no tiene lazo.

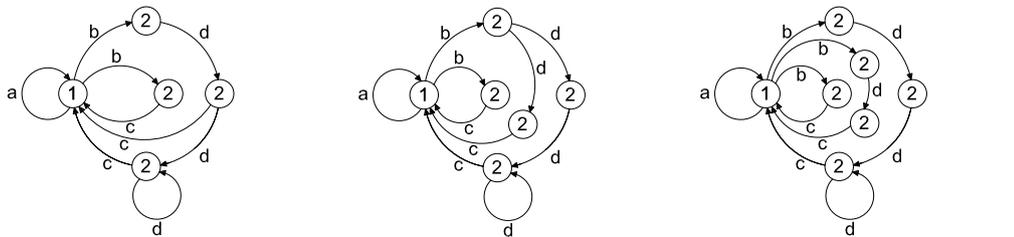


Figura 4.4: Paso 3, 4, 5 para generar el sistema de ciclos de primer retorno

Como resultado obtenemos la gráfica al centro de la Figura 4.4, a tal gráfica le aplicamos la operación de out-splitting al vértice que es vértice final la arista $\{b\}$ y vértice inicial de las aristas $\{d, d\}$. Como resultado obtenemos la gráfica del lado derecho de la Figura 4.4. Obsérvese que tenemos los ciclos de longitud 2,3 respectivamente, estos ya no se les aplica ninguna operación.

De esta forma se puede seguir una sucesión infinita de operaciones de in-splitting y out-splitting, de forma tal que la gráfica resultante es el sistema de ciclos de primer retorno en el vértice 1. Por como se van haciendo los splittings es claro que por cada $i \in \mathbb{N}$ obtendremos un solo

ciclo de longitud i , de esta forma nuestro sistema de ciclos de primer retorno esta descrito por la serie (4.4).

✦

Como vimos en el Ejemplo 4.3, el invariante del árbol generador es invariante bajo las operaciones de in-splitting y out-splitting, es lógico preguntarse si el invariante del árbol generador original es el mismo que el de los sistemas de ciclos de primer retorno. Abordaremos esta pregunta en la siguiente sección.

4.1.1. Presentación del problema

Los SFT's son una clase de σ -espacios muy estudiada, por lo que se conocen diferentes tipos de invariantes bajo conjugación como son la entropía, los puntos periódicos, la función zeta, etc. Uno de estos invariantes es el invariante de árbol generador, cuya demostración de su invarianza se encuentra en el capítulo 3 Teorema 3.7.

Gracias al Teorema de Descomposición 1.45 y la presentación en bloques de un σ -espacio, podemos determinar cuándo los σ -espacios generados por dos gráficas distintas son conjugados (Lema 1). Es decir si podemos ir de una gráfica a otra a través de un numero finito de las operaciones de splitting o amalgamaciones, los σ -espacios generados sobre estas gráficas son *conjugados*.

Como vimos en la sección anterior, a partir de una gráfica finita y un vértice en ésta, podemos construir el sistema de ciclos de primer retorno utilizando un número numerable de splittings. Por lo que la pregunta ¿habrá alguna relación entre el σ -espacio construido sobre la gráfica original y los σ -espacios construidos sobre los sistemas de ciclos de primer retorno? es natural. Como respuesta se sabe que el σ -espacio sobre la gráfica original y el σ -espacio sobre cualquiera de los sistemas de ciclos de primer retorno son en “esencia” los mismos. Es decir que son el mismo σ -espacio salvo un conjunto de medida 0.

Para de terminar si dos σ -espacios son en “esencia” los mismos necesitamos funciones entre los σ -espacios, las cuales tienen que conmutar con σ , ser medibles y ser en “esencia” biyectivas, los *isomorfismos*. Se sabe que algunos de los invariantes de conjugación mencionados con anterioridad siguen siendo invariantes bajo isomorfismos, funciones medibles que conmutan con la función σ y que son invertibles (casi donde sea), que *no son códigos de bloques*. Por la construcción del invariante del árbol generador y su demostración, nos hizo suponer que éste podría ser un invariante bajo las funciones que van de un σ -espacio generado por una gráfica G a un σ -espacio generado sobre los sistemas de ciclos de primer retorno de la gráfica G . Estas funciones forman parte de una clase de isomorfismos que se conocen como “isomorfismos de palabra mágica”, pero como se verá, el invariante del árbol generador no es un invariante bajo estas funciones. Los isomorfismos de “palabra mágica” son una clase de isomorfismos finitarios (que tienen una esperanza de tiempo de

codificación finito). En particular los isomorfismos de “palabra mágica”, son lo que nos hacen que el σ -espacio de Markov original sea “casi” igual al Markov shift generado por los sistemas de ciclos de primer retorno (véase [3]). Los isomorfismos finitarios tienen como característica principal que son continuas en casi todos lados, entendiéndose por ésto que el conjunto de puntos donde la función no es continua es de medida 0 bajo la medida establecida por la cadena de Markov.

Para una mejor exposición definamos qué es un isomorfismo de palabra mágica en el contexto de este trabajo.

Definición 4.5 (Palabra Mágica). *Sea $\varphi: X \rightarrow Y$ una función entre dos σ -espacios de Markov (son aquellos con entropía máxima). w una X -palabra (de tamaño $|w|$) es una palabra mágica si pasa lo siguiente:*

Para casi todas $x, x' \in X$, si $x[i, i + 2|w| + |\gamma| - 1] = w\gamma w = x'[i, i + 2|w| + |\gamma| - 1]$ para alguna $\gamma \in B(X)$ tal que $w\gamma w \in B(X)$ e $i \in \mathbb{Z}$, entonces $\varphi(x)[i + |w|; |\gamma| - 1] = \varphi(x')[i + |w|; |\gamma| - 1]$.

La siguiente es una definición equivalente.

Definición 4.6 (Palabra Mágica Inversa). *Sea $\varphi: X \rightarrow Y$ una función entre dos σ -espacios de Markov (son aquellos con entropía máxima). w una Y -palabra (de tamaño $|w|$) es una palabra mágica si pasa lo siguiente:*

1. *Si $y \in Y$ y $\{n \in \mathbb{Z} \mid y_{[n, n + |w| - 1]} = w\}$ no está acotado ni superior ni inferiormente, entonces y tiene una preimagen bajo φ .*
2. *Hay un entero I tal que siempre que u es una Y -palabra y dos puntos $x, x' \in X$ satisfacen $\varphi(x)_{[0, 2|w| + |u| - 1]} = wuw = \varphi(x')_{[0, 2|w| + |u| - 1]}$, entonces $x_{[I, I + |w| + |u| - 1]} = x'_{[I, I + |w| + |u| - 1]}$.*

Se sigue de (2) que la preimagen en (1) es única.

Definición 4.7 (Isomorfismo de Palabra Mágica). *Para X, Y σ -espacios de Markov (que son los σ -espacios con entropía máxima) con medidas invariantes $\mu \rightarrow \nu$, un isomorfismo de palabra mágica es un isomorfismo finitario $\varphi: (X, \mu) \rightarrow (Y, \nu)$ tal que ambas funciones φ y φ^{-1} tienen palabra mágica.*

Veamos que podemos definir un isomorfismo de palabra mágica entre el σ -espacio generado sobre una gráfica G y el σ -espacio generado sobre un sistema de ciclos de primer retorno. Sea G la gráfica original y H_v el sistema de ciclos de primer retorno en el vértice v . Vamos a definir $\phi: X_G \setminus M \rightarrow Y_{H_v} \setminus N$ donde M, N son conjuntos de medida 0. Para cualquier γ ciclo de primer retorno a v en la gráfica G , definimos a $\Phi(\gamma) = \bar{\gamma}$ y hacemos $\phi = \Phi_\infty$.

Primero observemos que ϕ no está definido para todos los puntos de X_G . Es decir, si $x \in X_G$ no visita nunca a v , o visita v un número finito de veces o deja de visitar v eventualmente, entonces $\phi(x)$ no está definida. El conjunto de puntos con dichas características es el conjunto M (el conjunto M es un subconjunto del complemento de los *puntos transitivos*, es decir, los puntos que visitan cualquier bloque una infinidad de veces en el pasado y en el futuro, y es conocido en el área que los puntos doblemente transitivos es un conjunto de medida 1 para cualquier medida de Markov ergódica).

Veamos que ϕ es un isomorfismo de palabra mágica, si ϕ es un isomorfismo de palabra mágica debería tener una palabra mágica: propongamos a v como tal (vamos a usar la Definición 4.5). Obsérvese que v cumple con ser una palabra mágica, es decir que para cualesquiera $x, x' \in X \setminus M$ tales que $x_{[i, i+2|v|+|u|-1]} = vuv = x'_{[i, i+2|v|+|u|-1]}$ con $u \in \mathcal{B}(X)$ tenemos que $\phi(x)$ y $\phi(x')$ están definidas pues los conjuntos $C = \{n \in \mathbb{Z} \mid x_{[n, n+|v|-1]} = v\}$ y $C' = \{n \in \mathbb{Z} \mid x'_{[n, n+|v|-1]} = v\}$ no están acotados, pues en $X \setminus M$ todos los elementos visitan una infinidad de veces v , tanto en el pasado como en el futuro. Como $vuv = \alpha$ es un ciclo de primer retorno a v , entonces al aplicar ϕ a x y x' tenemos que

$$\phi(x)_{[i+|v|, i+|\alpha|-1]} = \Phi(x_{[i, i+2|v|+|u|-1]}) = \Phi(\alpha) = \bar{\alpha} = \Phi(\alpha) = \Phi(x'_{[i, i+2|v|+|u|-1]}) = \phi(x')_{[i+|v|, i+|\alpha|-1]}$$

de esta forma v es una palabra mágica.

Ahora veamos que ϕ es un isomorfismo de palabra mágica, de la Definición 4.7 nos faltaría demostrar que ϕ es en “esencia” biyectiva y que ϕ^{-1} también tiene una palabra mágica. Por lo que veamos quién es ϕ^{-1} , sea $y \in \phi(X \setminus M)$ ¿quién es $\phi^{-1}(y)$? Como $y \in \phi(X \setminus M)$ se ve de la siguiente forma:

$$y = \dots \bar{\gamma}_{-2} \bar{\gamma}_{-1} \bar{\gamma}_0 \bar{\gamma}_1 \bar{\gamma}_2 \bar{\gamma}_3 \dots$$

donde $\bar{\gamma}_i$ es un ciclo en el sistema de ciclos de primer retorno. Hagamos notar que $\bar{\gamma}_i$ no es un solo símbolo o número, que $|\bar{\gamma}_i|$ y $|\bar{\gamma}_j|$ con $i \neq j$ no tienen por que ser iguales y $\bar{\gamma}_i \cap \bar{\gamma}_{i+1} = \bar{v}$. Ahora bien

$$\phi^{-1}(y) = \dots \gamma_{-2} \gamma_{-1} \gamma_0 \gamma_1 \gamma_2 \gamma_3 \dots$$

De esta forma al igual que la anterior notamos que $\gamma_i \cap \gamma_{i+1} = v$. A diferencia de ϕ veamos que para determinar la coordenada i en $\phi^{-1}(y)$ no necesitamos ver en donde se encuentra \bar{v} hacia el pasado o el futuro, pues con sólo saber en que vértice estamos se le asigna el valor correspondiente (ya sabemos en qu ciclo de primer retorno estamos al igual que el vrtice del ciclo en el que estamos). Es decir ϕ^{-1} es un 1-bloque.

El hecho que en “esencia” son iguales se da por que en un sistema de ciclos de primer retorno los puntos doblemente transitivos son los de medida 1 vease [3, 9]. De esta forma tenemos que entre el σ -espacio generado sobre una gráfica G y los σ -espacios generados sobre los sistemas de ciclos de primer retorno siempre hay un isomorfismo de palabra mágica.

Veamos que el Markov shift generado sobre un sistema de ciclos de primer retorno sigue siendo

un Markov shift fuertemente positivo recurrente, con ayuda del siguiente teorema que tomamos de [3] y que enunciamos en forma adecuada a nuestras necesidades..

Teorema 4.8. *Sean X y Y Markov shifts y supongamos X es fuertemente recurrente positivo con entropía finita. Si existe un isomorfismo de palabra mágica $\phi: X \rightarrow Y$ tal que ϕ^{-1} es un es 1-bloque, entonces Y es fuertemente recurrente positivo.*

La demostración se encuentra en [3].

Por el Teorema anterior y las observaciones hechas en esta sección podemos concluir que los σ -espacios sobre los sistemas de ciclos de primer retorno son fuertemente recurrentes positivos, por lo que no nos salimos de la convención de únicamente considerar cadenas de Markov fuertemente recurrentes positivas hecha en el Capítulo 2.

Aunque la suposición de que el invariante del árbol generador es invariante bajo isomorfismos de palabra mágica es incorrecta, se define el “invariante” del árbol generador para sistemas de ciclos de primer retorno. Se ve bajo qué condiciones se tienen el mismo valor para los diferentes sistemas de ciclos de primer retorno con base en los distintos vértices y las implicaciones con respecto a la matriz de la cadena de Markov asociada.

4.2. El invariante del árbol generador en los sistemas de ciclos de primer retorno

Pensando en la definición del invariante del árbol generador (véase [13] o la Definición 3.3), de igual forma vamos a definir el “invariante” del árbol generador en los sistemas de ciclos de primer retorno. El hecho de que el sistema de ciclos de primer retorno sea una gráfica infinita no afecta a la definición del invariante del árbol generador pues las expresiones (3.2), (3.4) y (3.5) tienen sentido como límites..

Veamos cuánto vale el “invariante” del árbol generador en los sistemas de ciclos de primer retorno. Para esto, sea P una matriz estocástica y sea $G = (V, E)$ la gráfica con matriz de adyacencia $P^\#$. Fijemos un vértice $v \in V$ y consideremos el sistema de ciclos de primer retorno a v descrito por la serie de potencias $f(t)$ (ver ecuaciones (4.1) y (4.2)). En la gráfica H que representa al sistema de ciclos de primer retorno en v , hay dos clases de árboles generadores. Si la raíz es \bar{v} , entonces es claro que sólo hay un árbol generador que resulta de quitar todas las flechas que salen de \bar{v} . Como los ciclos no se intersectan salvo en \bar{v} estos forman un árbol que contiene a todos los vértices de nuestro sistemas de ciclos de primer retorno. Por lo que sólo hay un árbol generador con raíz \bar{v} cuyo peso es 1.

Ahora el caso para los otros vértices. Sea $r \neq \bar{v}$ un vértice que está en el ciclo C_j^i , el j -ésimo

ciclo de longitud i . Es claro que el único árbol generador con raíz r resulta de eliminar la única arista con vértice inicial w y las aristas con vértice inicial \bar{v} excepto la que forma parte de C_j^i . Por lo que el árbol generador con raíz en \bar{v} tiene peso p_j^i .

Observación 4.9. *Es claro que todos los vértices en C_j^i excepto \bar{v} tienen el mismo peso para su árbol generador, por lo que si el ciclo es de longitud i entonces tenemos $i - 1$ árboles generadores cuyo peso es el mismo, por lo que el ciclo C_j^i aporta $(i - 1)p_j^i$.*

Entonces el invariante de árbol generador del sistema de ciclos de primer retorno a v queda de la siguiente forma:

$$1 + p_1^2 + p_2^2 \cdots + p_{K_2}^2 + 2p_1^3 + 2p_2^3 \cdots + 2p_{K_3}^3 + \cdots + (i - 1)p_1^i + \cdots + (i - 1)p_{K_i}^i + \cdots$$

$$1 + \underbrace{p_1^2 + p_2^2 \cdots + p_{K_2}^2}_{a_2} + 2 \underbrace{(p_1^3 + p_2^3 \cdots + p_{K_3}^3)}_{a_3} \cdots + (i - 1) \underbrace{(p_1^i + \cdots + p_{K_i}^i)}_{a_i} + \cdots$$

es decir,

$$1 + a_2 + 2a_3 + \cdots + (i - 1)a_i.$$

Esto es lo que vale nuestro invariante del árbol generador, en los sistemas de ciclos de primer retorno.

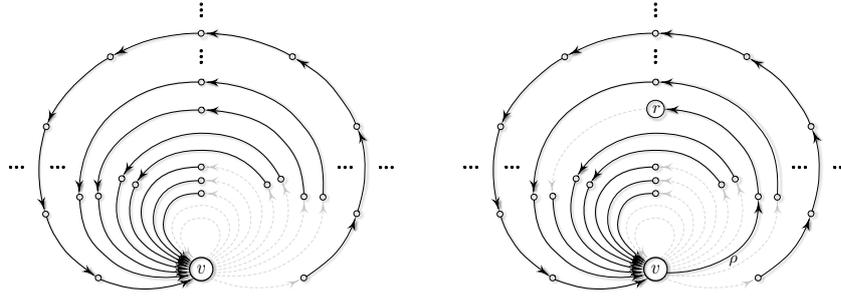


Figura 4.5: Tipos de árboles generadores en "Loop systems".

Enunciemos lo anteriormente probado en términos de la serie de potencias dada por (4.2).

Lema 4. *Sea P una matriz estocástica y sea $G = (V, E)$ la gráfica con matriz de adyacencia $P^\#$. Sea $v \in V$ y consideremos el sistema de ciclos de primer retorno en v dado por*

$$f_v(t) = \sum_{n=1}^{\infty} \sum_{\rho \in [0,1]} a_n^{(\rho)} [\rho] t^n.$$

Sea \hat{P} la matriz de transición de la cadena de Markov en las aristas de gráfica H que representa a los sistemas de ciclos de primer retorno en v . Entonces

$$\tau(\hat{P}) = \tau(f_v) \sum_{n=2}^{\infty} (n - 1) \sum_{\rho \in [0,1]} a_n^{(\rho)} \rho. \quad (4.10)$$

La fórmula 4.10 es una forma explícita del valor del invariante del árbol generador de un sistema de ciclos de primer retorno. El problema radica en “conocer” explícitamente la serie de potencias f . Como veremos, determinar la serie de potencias de un sistema de ciclos de primer retorno no es tarea fácil. En las siguientes secciones abordaremos este problema únicamente para matrices de tamaño 2 y 3. Primero calcularemos el invariante de árbol generador para los sistemas de ciclos de primer retorno de una cadena de Markov con dos estados.

4.2.1. El “2-shift”

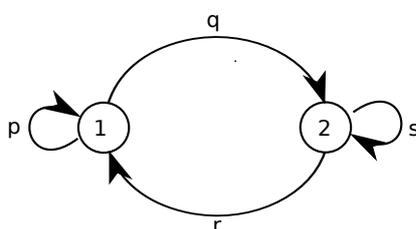


Figura 4.6: El 2-shift

La matriz de transición del 2-shifts de la figura 4.6 es la siguiente:

$$P = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$$

donde $p + q = 1$ y $r + s = 1$. Calculemos el sistema de ciclos de primer retorno en cualquiera de los 2 vértices.

Para el vértice 1 tenemos

$$\begin{aligned} 1 - f(t) &= \frac{\det(I - tP)}{1 - ts} \\ 1 - f(t) &= \frac{\begin{vmatrix} 1 - tp & -tq \\ -tr & 1 - st \end{vmatrix}}{1 - st} \\ 1 - f(t) &= \frac{1 - (s + p)t - (rq - ps)t^2}{1 - st} \\ 1 - f(t) &= (1 - st - st - spt^2 + rqt^2) \frac{1}{1 - st}. \end{aligned}$$

Expandiendo la parte de abajo de nuestra fracción nos queda

$$1 - f(t) = (1 - st - st - spt^2 + rqt^2)(1 + st + s^2t^2 + s^3t^3 + s^4t^4 \dots).$$

Multipliquemos lo que se encuentra dentro de los paréntesis

$$\begin{array}{cccccc}
 & 1 & +st & +s^2t^2 & +s^3t^3 & +s^4t^4 & \dots \\
 & & -st & -s^2t^2 & -s^3t^3 & -s^4t^4 & \dots \\
 1 - ft = & & -pt & -spt^2 & -s^2pt^3 & -s^3pt^4 & \dots \\
 & & & +spt^2 & +s^2pt^3 & +s^3pt^4 & \dots \\
 & & & -rpt^2 & -rqst^3 & -rs^2qt^4 & \dots \\
 \\
 1 - f(t) = & 1 & -pt & -rqt^2 & -rqst^3 & -rs^2qt^4 & \dots & rs^{i-2}qt^i \dots
 \end{array}$$

Entonces

$$f(t) = pt + rqt^2 + rqst^3 + rs^2qt^4 + \dots + rs^{i-2}qt^i \dots$$

Apliquemos la fórmula del Lema 4.10 para encontrar el invariante del árbol generador

$$\begin{aligned}
 & 1 + rq + 2rqs + 3rs^2q + 4rs^3q + \dots + irs^{i-1}q \\
 & 1 + rq \left(\sum_{i=1}^{\infty} is^{i-1} \right)
 \end{aligned}$$

utilicemos algo de cálculo

$$1 + rq \left(\sum_{i=1}^{\infty} \frac{dx^i}{dx} \Big|_s \right) = 1 + rq \left(\frac{d(\sum_{i=1}^{\infty} x^i)}{dx} \Big|_s \right) = 1 + rq \frac{d}{dx} \left(\frac{x}{1-x} \right) \Big|_s =$$

$$1 + rq \left(\frac{1}{(1-x)^2} \Big|_s \right) = 1 + \frac{rq}{(1-s)^2}.$$

Como $s + r = 1$ lo anterior nos queda como

$$1 + \frac{q}{r}.$$

Haciendo el mismo ejercicio para el sistema de ciclos de primer retorno del vértice 2 obtenemos que su invariante de árbol generador es

$$1 + \frac{r}{q}.$$

Acabamos de demostrar gran parte del siguiente teorema.

Teorema 4.11. *Sea*

$$P = \begin{pmatrix} p & q \\ r & s \end{pmatrix}.$$

Entonces

$$\tau(f_v) = \frac{\tau(P)}{\tau_v(P)} \tag{4.12}$$

para $v = 1, 2$. Por lo tanto, $\tau(f_1) = \tau(f_2)$ si y sólo si P es circulante.

En este caso, $\tau(f_1) = \tau(f_2) = 2$.

PRUEBA. Los invariantes del árbol generador de los sistemas de ciclos de primer retorno son iguales si y solo si

$$1 + \frac{q}{r} = 1 + \frac{r}{q}$$

y claramente esto ocurre si y solo si $q = r$. De forma que si los invariantes de árbol generador son iguales para los dos vértices, entonces la matriz de probabilidades de transición es de la siguiente forma.

$$\begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix}.$$

En este caso la matriz es circulante, y claramente $\tau(f_1) = \tau(f_2) = 2$, el valor Perron de la matriz de adyacencia del 2-shift. ✦

Vamos a concluir esta sección haciendo explícita nuestra técnica que usaremos para determinar los coeficientes de las series de potencias que presentan a los sistemas de ciclos de primer retorno: **DIAGONALIZACIÓN**. Usamos diagonalización para poder calcular para toda n las entradas de la matriz $(A^{(u)})^n$ en la ecuación (1.27). En el caso anterior de matrices de 2×2 , las matrices $A^{(u)}$ son de tamaño uno y no hubo necesidad de diagonalizar. Consideremos ahora matrices de tamaño 3.

4.2.2. El “3-shift”

En esta sección vamos a probar el teorema, que es el que corresponde al Teorema 4.11, para matrices de 3×3 . Veremos que la condición de ser circulante no es la que determina la igualdad de los invariantes del árbol generador de los sistemas de ciclos de primer retorno.

Teorema 4.13. *Sea*

$$P = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

Entonces

$$\tau(f_v) = \frac{\tau(P)}{\tau_v(P)} \tag{4.14}$$

para $v = 1, 2, 3$. Por lo tanto, $\tau(f_1) = \tau(f_2) = \tau(f_3)$ si y sólo si P es doblemente estocástica.

En este caso, $\tau(f_1) = \tau(f_2) = \tau(f_3) = 3$.

PRUEBA. Tomemos el 3-shift y calculemos los sistemas de ciclos de primer retorno. Para poder hacer tal cuenta observemos la matriz de probabilidades de transición

$$P = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

donde cumple la siguientes condiciones. $\{a, b, c, d, e, f, g, h, i\} \subset (0, 1)$ y $a + b + c = 1$, $d + e + f = 1$, $g + h + i = 1$. Es decir la matriz es una matriz estocástica, determinemos el sistema de ciclos de primer retorno utilizando la forma de teoría de gráficas.

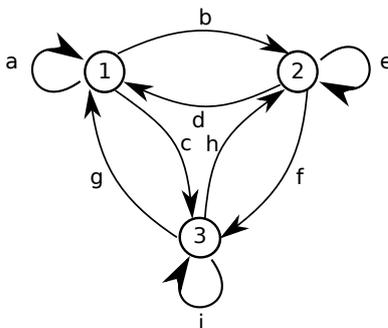


Figura 4.7: El 3-shift

Calculemos el sistema de ciclos primer retorno para el vértice 1. Es claro de la Figura 4.7 que sólo hay un camino de longitud uno, el cual es a ; de longitud 2 tenemos los caminos que nos da salir por la arista b y regresar por la arista d y de igual forma salir por la c y regresar por la g por lo que tenemos los caminos db y cg .

Ahora para todos los demás caminos hagamos una breve pausa y pongamos nuestra atención en la gráfica. Para cualquier camino de longitud n con $n \geq 3$ tenemos que salir del vértice 1 por lo que sólo podemos utilizar las aristas b o c y para regresar tenemos las aristas d o bien g . Podemos hacer un camino de longitud $n - 2$ en la subgráfica inducida por los vértices 2 y 3, y a este pegarle las aristas iniciales y finales generando así un camino de primer retorno a 1 de longitud n , de forma tal que dependiendo del vértice inicial le asignaremos la arista correspondiente, si el vértice inicial es el 2 tendrá la arista inicial b , si el vértice inicial es el 3 le corresponderá la arista inicial c . De igual forma asignaremos la arista final del “loop” de longitud n , d si el vértice final es el 2 o g si el vértice final es el 3.

Así los coeficientes de la serie de potencias del sistema de ciclos primer retorno del vértice 1 a partir de t^n con $n \geq 3$ están determinados por los caminos de longitud $n - 2$ en la gráfica inducida por los vértices 2 y 3.

Entonces para fijarnos en tales caminos nos fijamos en la matriz de adyacencia de tal gráfica inducida por 2 y 3

$$Q1 = \begin{pmatrix} e & f \\ h & i \end{pmatrix}.$$

Tal matriz es un menor de la matriz P que resulta de quitar el primer renglón y la primera columna. Para los caminos de longitud $n - 2$ que empiezan y terminan en el vértice 2 tenemos que dichos caminos son los de la entrada $Q1_{(1,1)}^{n-2}$, los que empiezan en 2 y terminan en 3 son la entrada $Q1_{(1,2)}^{n-2}$, los que empiezan en 3 y terminan en 2 son la entrada $Q1_{(2,1)}^{n-2}$ y finalmente los que empiezan en 3 y terminan en 3 están dados por $Q1_{(2,2)}^{n-2}$.

Una forma de ver lo anterior es notar que para calcular el sistema de ciclos primer retorno en realidad nos importa cuando regresan al vértice 1, es decir que si lo vemos como una caminata en la gráfica uno puede salir del vértice 1 y caminar por el resto de la gráfica pero al tomar las aristas de regreso al vértice 1 es cuando en realidad estamos contando lo que pasa.

Entonces nuestro sistema de ciclos primer retorno queda de la siguiente manera (ver (1.27)):

$$f_1(t) = at + (bd + cg)t^2 + \sum_{n=1}^{\infty} \left(bQ1_{(1,1)}^n d + bQ1_{(1,2)}^n g + cQ1_{(2,1)}^n d + cQ1_{(1,1)}^n g \right) t^{n+2}.$$

Para poder determinar los sistemas de ciclos primer retorno de los vértices 2 y 3 modificamos un poco lo anterior. Para el vértice 2 nos queda de la siguiente forma

$$Q2 = \begin{pmatrix} a & c \\ g & i \end{pmatrix},$$

análogamente para 3

$$Q3 = \begin{pmatrix} a & b \\ d & e \end{pmatrix}.$$

Ambas matrices resultan de quitar el correspondiente i -ésimo renglón y i -ésima columna. Por lo que los sistemas de ciclos primer retorno quedan de la siguiente forma

$$f_2(t) = et + (db + fh)t^2 + \sum_{n=1}^{\infty} \left(dQ2_{(1,1)}^n b + dQ2_{(1,2)}^n h + fQ2_{(2,1)}^n b + fQ2_{(2,2)}^n h \right) t^{n+2};$$

$$f_3(t) = it + (cg + fh)t^2 + \sum_{i=1}^{\infty} \left(gQ3_{(1,1)}^n c + gQ3_{(1,2)}^n f + gQ3_{(2,1)}^n f + hQ3_{(2,2)}^n f \right) t^{n+2}.$$

Observemos que si consideramos a $Qi^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ entonces podemos iniciar la serie desde 0. Esto nos ayuda en las cuentas pues tendremos un tipo de serie geométrica.

$$f_1(t) = at + \sum_{n=0}^{\infty} \left(bQ1_{(1,1)}^n d + bQ1_{(1,2)}^n g + cQ1_{(2,1)}^n d + cQ1_{(1,1)}^n g \right) t^{n+2};$$

$$f_2(t) = et + \sum_{n=0}^{\infty} \left(dQ2_{(1,1)}^n b + dQ2_{(1,2)}^n h + fQ2_{(2,1)}^n b + fQ2_{(2,2)}^n h \right) t^{n+2};$$

$$f_3(t) = it + \sum_{i=0}^{\infty} \left(gQ3_{(1,1)}^n c + gQ3_{(1,2)}^n f + hQ3_{(2,1)}^n c + hQ3_{(2,2)}^n f \right) t^{n+2}.$$

Ahora si aplicamos la fórmula del invariante del árbol generador a dichos sistemas de ciclos primer retorno obtenemos

$$\tau(f_1(t)) = 1 + \sum_{n=0}^{\infty} (n+1) \left(bQ1_{(1,1)}^n d + bQ1_{(1,2)}^n g + cQ1_{(2,1)}^n d + cQ1_{(1,1)}^n g \right);$$

$$\tau(f_2(t)) = 1 + \sum_{n=0}^{\infty} (n+1) \left(dQ2_{(1,1)}^n b + dQ2_{(1,2)}^n h + fQ2_{(2,1)}^n b + fQ2_{(2,2)}^n h \right);$$

$$\tau(f_3(t)) = 1 + \sum_{i=1}^{\infty} (n+1) \left(gQ3_{(1,1)}^n c + gQ3_{(1,2)}^n f + hQ3_{(2,1)}^n c + hQ3_{(2,2)}^n f \right).$$

Es claro de las fórmulas anteriores que para calcular el invariante del árbol generador necesitamos poder determinar las entradas de las matrices Q_i^n para $i \in \{1, 2, 3\}$ y $\forall n \in \mathbb{N}$.

Los coeficientes de $(Q_i)^n$. Tomemos como ejemplo a $Q1$, para poder determinar los coeficientes de $(Q1)^n$, nos ponemos a elevar a las diferentes potencias enteras a $Q1$.

$$Q1 = \begin{pmatrix} e & f \\ h & i \end{pmatrix}$$

$$(Q1)^2 = \begin{pmatrix} e^2 + fh & ef + fi \\ he + ih & fh + i^2 \end{pmatrix}$$

$$(Q1)^3 = \begin{pmatrix} e^3 + 2efh + hfi & fe^2 + f^2h + ief + fi^2 \\ he^2 + eih + fh^2 + hi^2 & efh + 2hfi + i^3 \end{pmatrix}$$

$$(Q1)^4 = \begin{pmatrix} e^4 + 3e^2fh + f^2h^2 + 2efih + fi^2h & e^3f + e^2fi + 2f^2he + 2f^2hi + efi^2 + fi^3 \\ he^3 + 2h^2ef + ihe^2 + 2ih^2f + i^2he + i^3h & e^2fh + 2efih + 3fi^2h + f^2h^2 + i^4 \end{pmatrix}$$

⋮

Es claro que intentar sacar una fórmula así nos llevará algo de tiempo, por lo cual recordando nuestras clases de Álgebra Lineal sabemos que si encontramos una matriz diagonal D similar a la matriz $Q1$, y encontrando una matriz B no singular entonces $Q1 = BDB^{-1}$ y tenemos que $(Q1)^n = BD^nB^{-1}$.

Para determinar las matrices diagonales que necesitamos, encontremos primero el valor característico de nuestras matrices Qi con $i \in \{1, 2, 3\}$.

Para $Q1 = \begin{pmatrix} e & f \\ h & i \end{pmatrix}$ entonces tomando el valor característico nos queda de la siguiente forma:

$$x_1 = \frac{1}{2}(e + i - \sqrt{\alpha}) \quad \text{y} \quad y_1 = \frac{1}{2}(e + i + \sqrt{\alpha})$$

en donde $\alpha = e^2 + 4fh - 2ei + i^2$.

Tomando los vectores característicos para cada raíz y colocándolos como columnas tenemos las siguientes matrices

$$B = \begin{pmatrix} \frac{e-i-\sqrt{\alpha}}{2h} & \frac{e-i+\sqrt{\alpha}}{2h} \\ 1 & 1 \end{pmatrix} \quad \text{y} \quad B^{-1} = \begin{pmatrix} -\frac{h}{\sqrt{\alpha}} & \frac{e-i+\sqrt{\alpha}}{2\sqrt{\alpha}} \\ \frac{h}{\sqrt{\alpha}} & -\frac{e-i-\sqrt{\alpha}}{2\sqrt{\alpha}} \end{pmatrix}.$$

Si D es una matriz diagonal con entradas x_1 y y_1 tenemos que $Q1^n = BD^nB^{-1}$. La matriz que deseamos nos queda de la siguiente forma

$$Q1^n = \begin{pmatrix} -\frac{e-i-\sqrt{\alpha}}{2\sqrt{\alpha}}x_1^n + \frac{e-i+\sqrt{\alpha}}{2\sqrt{\alpha}}y_1^n & -\frac{f}{\sqrt{\alpha}}x_1^n + \frac{f}{\sqrt{\alpha}}y_1^n \\ -\frac{h}{\sqrt{\alpha}}x_1^n + \frac{h}{\sqrt{\alpha}}y_1^n & \frac{e-i+\sqrt{\alpha}}{2\sqrt{\alpha}}x_1^n - \frac{e-i-\sqrt{\alpha}}{2\sqrt{\alpha}}y_1^n \end{pmatrix}.$$

Aplicando el mismo razonamiento las matrices $Q2$ y $Q3$ nos quedan como

$$Q2^n = \begin{pmatrix} -\frac{a-i-\sqrt{\beta}}{2\sqrt{\beta}}x_2^n + \frac{a-i+\sqrt{\beta}}{2\sqrt{\beta}}y_2^n & -\frac{c}{\sqrt{\beta}}x_2^n + \frac{c}{\sqrt{\beta}}y_2^n \\ -\frac{g}{\sqrt{\beta}}x_2^n + \frac{g}{\sqrt{\beta}}y_2^n & \frac{a-i-\sqrt{\beta}}{2\sqrt{\beta}}x_2^n - \frac{a-i+\sqrt{\beta}}{2\sqrt{\beta}}y_2^n \end{pmatrix};$$

$$Q3^n = \begin{pmatrix} -\frac{a-e-\sqrt{\gamma}}{2\sqrt{\gamma}}x_3^n + \frac{a-e+\sqrt{\gamma}}{2\sqrt{\gamma}}y_3^n & -\frac{b}{\sqrt{\gamma}}x_3^n + \frac{b}{\sqrt{\gamma}}y_3^n \\ -\frac{d}{\sqrt{\gamma}}x_3^n + \frac{d}{\sqrt{\gamma}}y_3^n & \frac{a-e+\sqrt{\gamma}}{2\sqrt{\gamma}}x_3^n - \frac{a-e-\sqrt{\gamma}}{2\sqrt{\gamma}}y_3^n \end{pmatrix};$$

siendo $x_2 = \frac{1}{2}(a + i - \sqrt{\beta})$, $y_2 = \frac{1}{2}(a + i + \sqrt{\beta})$, $x_3 = \frac{1}{2}(e + a - \sqrt{\gamma})$, $y_3 = \frac{1}{2}(e + a + \sqrt{\gamma})$ y $\beta = i^2 - 2ia + 4gc + a^2$, $\gamma = e^2 - 2ea + 4db + a^2$ respectivamente.

Sustituyendo los coeficientes de las Q_i en las fórmulas del invariante del árbol generador para los sistemas de ciclos primer retorno nos quedan como:

$$\tau(f_1(t)) = 1 + \sum_{n=0}^{\infty} (n+1) \left(b \left(-\frac{e-i-\sqrt{\alpha}}{2\sqrt{\alpha}} x_1^n + \frac{e-i+\sqrt{\alpha}}{2\sqrt{\alpha}} y_1^n \right) d + b \left(-\frac{f}{\sqrt{\alpha}} x_1^n + \frac{f}{\sqrt{\alpha}} y_1^n \right) g + \right. \\ \left. + c \left(-\frac{h}{\sqrt{\alpha}} x_1^n + \frac{h}{\sqrt{\alpha}} y_1^n \right) d + c \left(\frac{e-i+\sqrt{\alpha}}{2\sqrt{\alpha}} x_1^n - \frac{e-i-\sqrt{\alpha}}{2\sqrt{\alpha}} y_1^n \right) g \right) ;$$

$$\tau(f_2(t)) = 1 + \sum_{n=0}^{\infty} (n+1) \left(d \left(-\frac{a-i-\sqrt{\beta}}{2\sqrt{\beta}} x_2^n + \frac{a-i+\sqrt{\beta}}{2\sqrt{\beta}} y_2^n \right) b + d \left(-\frac{c}{\sqrt{\beta}} x_2^n + \frac{c}{\sqrt{\beta}} y_2^n \right) h + \right. \\ \left. + f \left(-\frac{g}{\sqrt{\beta}} x_2^n + \frac{g}{\sqrt{\beta}} y_2^n \right) b + f \left(\frac{a-i-\sqrt{\beta}}{2\sqrt{\beta}} x_2^n - \frac{a-i+\sqrt{\beta}}{2\sqrt{\beta}} y_2^n \right) h \right) ;$$

$$\tau(f_3(t)) = 1 + \sum_{i=0}^{\infty} (n+1) \left(g \left(-\frac{a-e-\sqrt{\gamma}}{2\sqrt{\gamma}} x_3^n + \frac{a-e+\sqrt{\gamma}}{2\sqrt{\gamma}} y_3^n \right) c + g \left(-\frac{b}{\sqrt{\gamma}} x_3^n + \frac{b}{\sqrt{\gamma}} y_3^n \right) f + \right. \\ \left. + h \left(-\frac{d}{\sqrt{\gamma}} x_3^n + \frac{d}{\sqrt{\gamma}} y_3^n \right) c + h \left(\frac{a-e+\sqrt{\gamma}}{2\sqrt{\gamma}} x_3^n - \frac{a-e-\sqrt{\gamma}}{2\sqrt{\gamma}} y_3^n \right) f \right) .$$

Simplificando tales ecuaciones y factorizando términos comunes las ecuaciones anteriores quedan como

$$\tau(f_1(t)) = 1 + \sum_{n=0}^{\infty} (n+1) \left(-\frac{b(e-i-\sqrt{\alpha})d}{2\sqrt{\alpha}} x_1^n - \frac{bfg}{\sqrt{\alpha}} x_1^n - \frac{chd}{\sqrt{\alpha}} x_1^n + \frac{c(e-i+\sqrt{\alpha})g}{2\sqrt{\alpha}} x_1^n + \right. \\ \left. + \frac{b(e-i+\sqrt{\alpha})d}{2\sqrt{\alpha}} y_1^n + \frac{bfg}{\sqrt{\alpha}} y_1^n + \frac{chd}{\sqrt{\alpha}} y_1^n - \frac{c(e-i-\sqrt{\alpha})g}{2\sqrt{\alpha}} y_1^n \right) ;$$

$$\tau(f_2(t)) = 1 + \sum_{n=0}^{\infty} (n+1) \left(-\frac{d(a-i-\sqrt{\beta})b}{2\sqrt{\beta}} x_2^n - \frac{dch}{\sqrt{\beta}} x_2^n - \frac{fgb}{\sqrt{\beta}} x_2^n + \frac{f(a-i-\sqrt{\beta})h}{2\sqrt{\beta}} x_2^n + \right. \\ \left. + \frac{d(a-i+\sqrt{\beta})b}{2\sqrt{\beta}} y_2^n + \frac{dch}{\sqrt{\beta}} y_2^n + \frac{fgb}{\sqrt{\beta}} y_2^n - \frac{f(a-i+\sqrt{\beta})h}{2\sqrt{\beta}} y_2^n \right) ;$$

$$\tau(f_3(t)) = 1 + \sum_{i=0}^{\infty} (n+1) \left(-\frac{g(a-e-\sqrt{\gamma})c}{2\sqrt{\gamma}} x_3^n - \frac{gbf}{\sqrt{\gamma}} x_3^n - \frac{hdc}{\sqrt{\gamma}} x_3^n + \frac{h(a-e+\sqrt{\gamma})f}{2\sqrt{\gamma}} x_3^n + \right. \\ \left. + \frac{g(a-e+\sqrt{\gamma})c}{2\sqrt{\gamma}} y_3^n + \frac{gbf}{\sqrt{\gamma}} y_3^n + \frac{hdc}{\sqrt{\gamma}} y_3^n - \frac{h(a-e-\sqrt{\gamma})f}{2\sqrt{\gamma}} y_3^n \right) ,$$

$$\tau(f_1(t)) = 1 + \sum_{n=0}^{\infty} (n+1) \left(\left(-\frac{b(e-i-\sqrt{\alpha})d}{2\sqrt{\alpha}} - \frac{bfg}{\sqrt{\alpha}} - \frac{chd}{\sqrt{\alpha}} + \frac{c(e-i+\sqrt{\alpha})g}{2\sqrt{\alpha}} \right) x_1^n + \right. \\ \left. + \left(\frac{b(e-i+\sqrt{\alpha})d}{2\sqrt{\alpha}} + \frac{bfg}{\sqrt{\alpha}} + \frac{chd}{\sqrt{\alpha}} - \frac{c(e-i-\sqrt{\alpha})g}{2\sqrt{\alpha}} \right) y_1^n \right) ;$$

$$\tau(f_2(t)) = 1 + \sum_{n=0}^{\infty} (n+1) \left(\left(-\frac{d(a-i-\sqrt{\beta})b}{2\sqrt{\beta}} - \frac{dch}{\sqrt{\beta}} - \frac{fgb}{\sqrt{\beta}} + \frac{f(a-i-\sqrt{\beta})h}{2\sqrt{\beta}} \right) x_2^n + \right. \\ \left. + \left(\frac{d(a-i+\sqrt{\beta})b}{2\sqrt{\beta}} + \frac{dch}{\sqrt{\beta}} + \frac{fgb}{\sqrt{\beta}} - \frac{f(a-i+\sqrt{\beta})h}{2\sqrt{\beta}} \right) y_2^n \right) ;$$

$$\tau(f_3(t)) = 1 + \sum_{i=0}^{\infty} (n+1) \left(\left(-\frac{g(a-e-\sqrt{\gamma})c}{2\sqrt{\gamma}} - \frac{gbf}{\sqrt{\gamma}} - \frac{hdc}{\sqrt{\gamma}} + \frac{h(a-e+\sqrt{\gamma})f}{2\sqrt{\gamma}} \right) x_3^n + \left(\frac{g(a-e+\sqrt{\gamma})c}{2\sqrt{\gamma}} + \frac{gbf}{\sqrt{\gamma}} + \frac{hdc}{\sqrt{\gamma}} - \frac{h(a-e-\sqrt{\gamma})f}{2\sqrt{\gamma}} \right) y_3^n \right).$$

Como se ve la ecuaciones empiezan a ser muy largas, por lo que haremos una sustitución para que las cuentas se puedan seguir con mayor facilidad

$$\begin{aligned} A_1 &= \left(-\frac{b(e-i-\sqrt{\alpha})d}{2\sqrt{\alpha}} - \frac{bfg}{\sqrt{\alpha}} - \frac{chd}{\sqrt{\alpha}} + \frac{c(e-i+\sqrt{\alpha})g}{2\sqrt{\alpha}} \right) \quad \text{y} \quad B_1 = \left(\frac{b(e-i+\sqrt{\alpha})d}{2\sqrt{\alpha}} + \frac{bfg}{\sqrt{\alpha}} + \frac{chd}{\sqrt{\alpha}} - \frac{c(e-i-\sqrt{\alpha})g}{2\sqrt{\alpha}} \right), \\ A_2 &= \left(-\frac{d(a-i-\sqrt{\beta})b}{2\sqrt{\beta}} - \frac{dch}{\sqrt{\beta}} - \frac{fjb}{\sqrt{\beta}} + \frac{f(a-i-\sqrt{\beta})h}{2\sqrt{\beta}} \right) \quad \text{y} \quad B_2 = \left(\frac{g(a-e+\sqrt{\gamma})c}{2\sqrt{\gamma}} + \frac{gbf}{\sqrt{\gamma}} + \frac{hdc}{\sqrt{\gamma}} - \frac{h(a-e-\sqrt{\gamma})f}{2\sqrt{\gamma}} \right), \\ A_3 &= \left(-\frac{g(a-e-\sqrt{\gamma})c}{2\sqrt{\gamma}} - \frac{gbf}{\sqrt{\gamma}} - \frac{hdc}{\sqrt{\gamma}} + \frac{h(a-e+\sqrt{\gamma})f}{2\sqrt{\gamma}} \right) \quad \text{y} \quad B_3 = \left(\frac{g(a-e+\sqrt{\gamma})c}{2\sqrt{\gamma}} + \frac{gbf}{\sqrt{\gamma}} + \frac{hdc}{\sqrt{\gamma}} - \frac{h(a-e-\sqrt{\gamma})f}{2\sqrt{\gamma}} \right). \end{aligned}$$

Observación 4.15. Para la cuenta será importante observar los valores de $A_i + B_i$ así como $A_i - B_i$ con $i \in \{1, 2, 3\}$

$$\begin{aligned} A_1 + B_1 &= bd + cg & A_1 - B_1 &= \frac{bdj - bde - 2bfg - 2chd - cgj + cge}{\sqrt{\alpha}}; \\ A_2 + B_2 &= db + fh & A_2 - B_2 &= \frac{dbj - dba - 2dch - 2fjb - fhj + fha}{\sqrt{\beta}}; \\ A_3 + B_3 &= gc + hf & A_3 - B_3 &= \frac{gce - gca - 2gbf - 2hdc - hfe + hfa}{\sqrt{\gamma}}. \end{aligned}$$

Con dichas sustituciones, los invariante del árbol generador para los sistemas de ciclos primer retorno nos quedan como

$$\begin{aligned} \tau(f_1(t)) &= 1 + \sum_{n=0}^{\infty} (A_1 x_1^n + B_1 y_1^n); \\ \tau(f_2(t)) &= 1 + \sum_{n=0}^{\infty} (A_2 x_2^n + B_2 y_2^n); \\ \tau(f_3(t)) &= 1 + \sum_{n=0}^{\infty} (A_3 x_3^n + B_3 y_3^n). \end{aligned}$$

Como $x_1, x_2, x_3, y_1, y_2, y_3 < 1$ tenemos una serie geométrica. Aplicando algo de cálculo tenemos

$$\begin{aligned} \tau(f_1(t)) &= 1 + \left(\frac{A_1}{(1-x_1)^2} + \frac{B_1}{(1-y_1)^2} \right); \\ \tau(f_2(t)) &= 1 + \left(\frac{A_2}{(1-x_2)^2} + \frac{B_2}{(1-y_2)^2} \right); \end{aligned}$$

$$\tau(f_3(t)) = 1 + \left(\frac{A_3}{(1-x_3)^2} + \frac{B_3}{(1-y_3)^2} \right).$$

Ya que tenemos una expresión más sencilla sustituycamos a x_i, y_i con $i \in \{1, 2, 3\}$ por su valores

$$\begin{aligned} \tau(f_1(t)) &= 1 + \left(\frac{A_1}{(1-\frac{1}{2}(e+i-\sqrt{\alpha}))^2} + \frac{B_1}{(1-\frac{1}{2}(e+i+\sqrt{\alpha}))^2} \right); \\ \tau(f_2(t)) &= 1 + \left(\frac{A_2}{(1-\frac{1}{2}(a+i-\sqrt{\beta}))^2} + \frac{B_2}{(1-\frac{1}{2}(a+i+\sqrt{\beta}))^2} \right); \\ \tau(f_3(t)) &= 1 + \left(\frac{A_3}{(1-\frac{1}{2}(e+a-\sqrt{\gamma}))^2} + \frac{B_3}{(1-\frac{1}{2}(e+a+\sqrt{\gamma}))^2} \right); \\ \tau(f_1(t)) &= 1 + \left(\frac{A_1(1-\frac{1}{2}(e+i+\sqrt{\alpha}))^2 + B_1(1-\frac{1}{2}(e+i-\sqrt{\alpha}))^2}{(1-\frac{1}{2}(e+i-\sqrt{\alpha}))^2(1-\frac{1}{2}(e+i+\sqrt{\alpha}))^2} \right); \\ \tau(f_2(t)) &= 1 + \left(\frac{A_2(1-\frac{1}{2}(a+i+\sqrt{\beta}))^2 + B_2(1-\frac{1}{2}(a+i-\sqrt{\beta}))^2}{(1-\frac{1}{2}(a+i-\sqrt{\beta}))^2(1-\frac{1}{2}(a+i+\sqrt{\beta}))^2} \right); \\ \tau(f_3(t)) &= 1 + \left(\frac{A_3(1-\frac{1}{2}(e+a+\sqrt{\gamma}))^2 + B_3(1-\frac{1}{2}(e+a-\sqrt{\gamma}))^2}{(1-\frac{1}{2}(e+a-\sqrt{\gamma}))^2(1-\frac{1}{2}(e+a+\sqrt{\gamma}))^2} \right); \\ \tau(f_1(t)) &= 1 + \left(\frac{4(A_1(-2+j+e+\sqrt{\alpha})^2 + B_1(-2+j+e+\sqrt{\alpha})^2)}{(-2+j+e-\sqrt{\alpha})^2(-2+j+e+\sqrt{\alpha})^2} \right); \\ \tau(f_2(t)) &= 1 + \left(\frac{4(A_2(-2+j+a-\sqrt{\beta})^2 + B_2(-2+j+a+\sqrt{\beta})^2)}{(-2+j+a-\sqrt{\beta})^2(-2+j+a+\sqrt{\beta})^2} \right); \\ \tau(f_3(t)) &= 1 + \left(\frac{4(A_3(-2+e+a-\sqrt{\gamma})^2 + B_3(-2+e+a+\sqrt{\gamma})^2)}{(-2+e+a-\sqrt{\gamma})^2(-2+e+a+\sqrt{\gamma})^2} \right). \end{aligned}$$

Enfoquemos nuestra atención en lo que está dentro de los paréntesis, expandiendo los términos elevados al cuadrado y factorizando el divisor, se obtiene

$$\begin{aligned} \tau(f_1(t)) &= 1 + \lambda 4(4A_1 - 4A_1j - 4A_1e - 4A_1\sqrt{\alpha} + A_1j^2 + 2A_1je + 2A_1j\sqrt{\alpha} + A_1e^2 + \\ &\quad + 2A_1e\sqrt{\alpha} + A_1\alpha + 4B_1 - 4B_1j - 4B_1e + 4B_1\sqrt{\alpha} + \\ &\quad + B_1j^2 + 2B_1je - 2B_1j\sqrt{\alpha} + B_1e^2 - 2B_1e\sqrt{\alpha} + B_1\alpha); \\ \tau(f_2(t)) &= 1 + \mu 4(4A_2 - 4A_2j - 4A_2a - 4A_2\sqrt{\beta} + A_2j^2 + 2A_2ja + 2A_2j\sqrt{\beta} + A_2a^2 + \\ &\quad + 2A_2a\sqrt{\beta} + A_2\beta + 4B_2 - 4B_2j - 4B_2a + 4B_2\sqrt{\beta} + \\ &\quad + B_2j^2 + 2B_2ja - 2B_2j\sqrt{\beta} + B_2a^2 - 2B_2a\sqrt{\beta} + A_2\beta); \end{aligned}$$

$$\begin{aligned}\tau(f_3(t)) = 1 + & \nu 4(4A_3 - 4A_3e - 4A_3a - 4A_3\sqrt{\gamma} + A_3e^2 + 2A_3ea + 2A_3e\sqrt{\gamma} + A_3a^2 + \\ & + 2A_3a\sqrt{\gamma} + A_3\gamma + 4B_3 - 4B_3e - 4B_3a + 4B_3\sqrt{\gamma} + \\ & + B_3e^2 + 2B_3ea - 2B_3e\sqrt{\gamma} + B_3a^2 - 2B_3a\sqrt{\gamma} + B_3\gamma); \end{aligned}$$

con

$$\begin{aligned}\lambda &= \frac{1}{(-2 + j + e - \sqrt{\alpha})^2(-2 + j + e + \sqrt{\alpha})^2} \\ \mu &= \frac{1}{(-2 + j + a - \sqrt{\beta})^2(-2 + j + a + \sqrt{\beta})^2} \\ \nu &= \frac{1}{(-2 + e + a - \sqrt{\gamma})^2(-2 + e + a + \sqrt{\gamma})^2}.\end{aligned}$$

Agrupemos como nos sugiere nuestra escritura de las ecuaciones

$$\begin{aligned}\tau(f_1(t)) = 1 + & \lambda 4(4(A_1 + B_1) - 4(A_1 + B_1)j - 4(A_1 + B_1)e - 4(A_1 - B_1)\sqrt{\alpha} + \\ & + (A_1 + B_1)j^2 + 2(A_1 + B_1)je + 2(A_1 - B_1)j\sqrt{\alpha} + \\ & + (A_1 + B_1)e^2 + 2(A_1 - B_1)e\sqrt{\alpha} + (A_1 + B_1)\alpha); \end{aligned}$$

$$\begin{aligned}\tau(f_2(t)) = 1 + & \mu 4(4(A_2 + B_2) - 4(A_2 + B_2)j - 4(A_2 + B_2)a - 4(A_2 - B_2)\sqrt{\beta} + \\ & + (A_2 + B_2)j^2 + 2(A_2 + B_2)ja + 2(A_2 - B_2)j\sqrt{\beta} + \\ & + (A_2 + B_2)a^2 + 2(A_2 - B_2)a\sqrt{\beta} + (A_2 + B_2)\beta); \end{aligned}$$

$$\begin{aligned}\tau(f_3(t)) = 1 + & \nu 4(4(A_3 + B_3) - 4(A_3 + B_3)e - 4(A_3 + B_3)a - 4(A_3 - B_3)\sqrt{\gamma} + \\ & + (A_3 + B_3)e^2 + 2(A_3 + B_3)ea + 2(A_3 - B_3)e\sqrt{\gamma} + \\ & + (A_3 + B_3)a^2 + 2(A_3 - B_3)a\sqrt{\gamma} + (A_3 + B_3)\gamma). \end{aligned}$$

Recordando la Observación 4.15 se tiene que

$$\begin{aligned}\tau(f_1(t)) = 1 + & \lambda 4(4(bd + c) - 4(bd + cg)j - 4(bd + cg)e + \\ & - 4 \left(\frac{bdj - bde - 2bfg - 2chd - cgj + cge}{\sqrt{\alpha}} \right) \sqrt{\alpha} + (bd + cg)j^2 + 2(bd + cg)je + \\ & + 2 \left(\frac{bdj - bde - 2bfg - 2chd - cgj + cge}{\sqrt{\alpha}} \right) j\sqrt{\alpha} + (bd + cg)e^2 + \\ & + 2 \left(\frac{bdj - bde - 2bfg - 2chd - cgj + cge}{\sqrt{\alpha}} \right) e\sqrt{\alpha} + (bd + cg)\alpha); \end{aligned}$$

$$\begin{aligned}\tau(f_2(t)) = 1 + & \mu 4(4(bd + fh) - 4(bd + fh)j - 4(bd + fh)a + \\ & - 4 \left(\frac{dbj - dba - 2dch - 2fgb - fhj + fha}{\sqrt{\beta}} \right) \sqrt{\beta} + (bd + fh)j^2 + 2(bd + fh)ja + \\ & + 2 \left(\frac{dbj - dba - 2dch - 2fgb - fhj + fha}{\sqrt{\beta}} \right) j\sqrt{\beta} + (bd + fh)a^2 + \\ & + 2 \left(\frac{dbj - dba - 2dch - 2fgb - fhj + fha}{\sqrt{\beta}} \right) a\sqrt{\beta} + (bd + fh)\beta); \end{aligned}$$

$$\begin{aligned}\tau(f_3(t)) = 1 + & \nu 4(4(gc + hf) - 4(gc + hf)e - 4(gc + hf)a + \\ & - 4 \left(\frac{gce - gca - 2gbf - 2hdc - hfe + hfa}{\sqrt{\gamma}} \right) \sqrt{\gamma} + (gc + hf)e^2 + 2(gc + hf)ea + \\ & + 2 \left(\frac{gce - gca - 2gbf - 2hdc - hfe + hfa}{\sqrt{\gamma}} \right) e\sqrt{\gamma} + (gc + hf)a^2 + \\ & + 2 \left(\frac{gce - gca - 2gbf - 2hdc - hfe + hfa}{\sqrt{\gamma}} \right) a\sqrt{\gamma} + (gc + hf)\gamma). \end{aligned}$$

Como se aprecia los coeficientes $\sqrt{\alpha}$, $\sqrt{\beta}$ y $\sqrt{\gamma}$ desaparecen del interior del paréntesis principal en las ecuaciones. Como la matriz es estocástica, sabemos que $a = 1 - b - c$, $e = 1 - d - f$ y $i = 1 - g - h$ sustituyendo nos quedan como

$$\tau(f_1(t)) = 1 + \lambda 4(4(fg + hd + gd)(fb + fc + bg + hb + hc + dc));$$

$$\tau(f_2(t)) = 1 + \mu 4(4(gb + hb + hc)(fb + fc + fg + dc + dg + dh));$$

$$\tau(f_3(t)) = 1 + \nu 4(4(dc + fb + fc)(gb + gd + gf + hb + hc + hd)).$$

Si en λ, μ, ν sustituimos α, β, γ respectivamente obtendremos que

$$\lambda = \frac{1}{16(gf+gd+hd)^2}, \quad \mu = \frac{1}{16(gb+hb+hc)^2}, \quad \nu = \frac{1}{16(dc+fb+fc)^2}.$$

Sustituyendo en nuestras fórmulas nos da

$$\tau(f_1(t)) = 1 + \frac{fb + fc + bg + hb + hc + dc}{fg + hd + gd}; \quad (4.16)$$

$$\tau(f_2(t)) = 1 + \frac{fb + fc + fg + dc + dg + dh}{gb + hb + hc}; \quad (4.17)$$

$$\tau(f_3(t)) = 1 + \frac{gb + gd + gf + hb + hc + hd}{dc + fb + fc}. \quad (4.18)$$

Observación 4.19. Tomemos $\tau(f_1(t))$ el invariante del árbol generador en el sistema de ciclos de primer retorno en el vértice 1. En el numerador de la fracción tenemos la suma de todos los posibles árboles generadores con raíz el vértice 2 y vértice 3, mientras que en el denominador tenemos a la suma de los pesos de todos los posibles árboles generadores con raíz vértice 1. Esto pasa para cualquiera de los tres vértices. Por lo que las ecuaciones (4.16), (4.17) y (4.18) las podemos ver como:

$$\tau(f_1(t)) = \frac{\tau(P)}{\tau_1(P)};$$

$$\tau(f_2(t)) = \frac{\tau(P)}{\tau_2(P)};$$

$$\tau(f_3(t)) = \frac{\tau(P)}{\tau_3(P)}.$$

Por lo tanto gracias a la observación anterior tenemos que se cumple la Ecuación (4.14).

Ahora que tenemos de forma explícita las fórmulas para los invariantes del árbol generador para los diferentes sistemas de ciclos de primer retorno, veamos que para las matrices circulantes cumplen que son iguales y tienen valor 3.

Matrices circulares de 3×3 . Tomemos una matriz de transición de probabilidades circulante

$$P = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}.$$

Es claro que en cualquiera de los 3 vértices tenemos que sus sistemas de ciclos de primer retorno es el mismo, por lo tanto sus invariantes del árbol generador tendrán el mismo valor. Calculemos dicho valor, para ésto utilicemos cualquiera de las Fórmulas (4.16), (4.17) y (4.18). Si sustituimos a d, e, f, g, h, i en (4.16) de la siguiente forma:

$$\begin{aligned} e &= i = a \\ g &= f = b \\ d &= h = c. \end{aligned}$$

Obtenemos que queda como

$$\tau(f_1(t)) = 1 + \frac{b^2 + bc + b^2 + cb + c^2 + c^2}{b^2 + bc + c^2} = 1 + \frac{2b^2 + 2bc + 2c^2}{b^2 + bc + c^2} = 1 + 2 \frac{b^2 + bc + c^2}{b^2 + bc + c^2} = 3.$$

Hagamos una pausa en la desmotrasción para considerar un ejemplo importante. Si la matriz de transición de probabilidades es una matriz circulante, entonces el invariante del árbol generador en los sistemas de ciclos primer retorno son iguales y tendrán valor 3. Sin embargo el regreso no es cierto, es decir, si los invariantes del árbol generador son iguales para los sistemas de ciclos de primer retorno para cada vértice, entonces la matriz de transición de probabilidades no será necesariamente una matriz circulante.

Ejemplo 4.20. Sea $\begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 2/6 & 3/6 & 1/6 \\ 2/6 & 1/6 & 3/6 \end{pmatrix}$. Es claro que tal matriz no es una matriz circulante, podemos calcular el invariante del árbol generador para cada uno de los sistemas de ciclos primer retorno. Como resultado tenemos que cada uno vale 3.

Como hemos demostrado si una matriz es circulante el invariante del árbol generador en los sistemas de ciclos de primer retorno en cada vértice vale lo mismo, con el ejemplo se ve que el regreso no es cierto. Entonces ¿qué condiciones tenemos que pedir a la matriz P para que los invariantes del árbol generador de los sistemas de ciclos de primer retorno valgan lo mismo y consideren que contengan al conjunto de matrices circulares? La respuesta a tal pregunta son las matrices doblemente estocásticas. Al respecto se han hecho muchos artículos pero en ninguno se ha encontrado una propiedad similar a ésta. Entonces para el 2-shift reformulemos el Teorema 4.11.

Teorema 4.21. *Sea*

$$P = \begin{pmatrix} p & q \\ r & s \end{pmatrix}.$$

Entonces

$$\tau(f_v) = \frac{\tau(P)}{\tau_v(P)} \quad (4.22)$$

para $v = 1, 2$. Por lo tanto, $\tau(f_1) = \tau(f_2)$ si y sólo si una de las siguientes afirmaciones se satisface.

1. P es circulante.
2. P es doblemente estocástica.

En este caso, $\tau(f_1) = \tau(f_2) = 2$.

Ahora continuemos con la demostación del Teorema 4.13. Ya que sabemos cuánto valen los invariantes del árbol generador para cada uno de los sistemas de ciclos primer retorno de nuestra gráfica la pregunta que nos habíamos inicialmente planteado se reduce a saber cuándo pasa que

$$\frac{fb + fc + bg + hb + hc + dc}{fg + hd + gd} = \frac{fb + fc + fg + dc + dg + dh}{gb + hb + hc} = \frac{gb + gd + gf + hb + hc + hd}{dc + fb + fc}.$$

Tomando la primera igualdad tenemos que

$$\frac{fb + fc + bg + hb + hc + dc}{fg + hd + gd} - \frac{fb + fc + fg + dc + dg + dh}{gb + hb + hc} = 0$$

realizado la resta y factorizando adecuadamente nos queda como

$$\frac{(bg - dg - fg + bh + ch - dh)(cd + bf + cf + bg + dg + fg + bh + ch + dh)}{(dg + dh + fg)(bg + bh + ch)}.$$

Como P es una matriz estocástica todos los elementos de ésta son mayores o iguales que 0, por lo tanto si la expresión está igualada a 0 sólo nos queda que $bg - dg - fg + bh + ch - dh = 0$.

De igual forma hacemos lo mismo para las otras igualdades:

$$\frac{fb + fc + bg + hb + hc + dc}{fg + hd + gd} - \frac{gb + gd + gf + hb + hc + hd}{dc + fb + fc} = 0,$$

entonces nos queda $cd + bf + cf - dg - fg - dh = 0$ y

$$\frac{fb + fc + fg + dc + dg + dh}{gb + hb + hc} - \frac{gb + gd + gf + hb + hc + hd}{dc + fb + fc} = 0$$

tenemos que $cd + bf + cf - bg - bh - ch = 0$.

Entonces si sabemos que los invariantes del árbol generador son el mismo para los tres distintos vértices tenemos el siguiente sistema de ecuaciones

$$bg - dg - fg + bh + ch - dh = 0. \quad (4.23)$$

$$cd + bf + cf - dg - fg - dh = 0. \quad (4.24)$$

$$cd + bf + cf - bg - bh - ch = 0. \quad (4.25)$$

Es claro que el sistema no tiene solución única, pero nos da condiciones sobre la matriz tomando las ecuaciones por pares. Empezemos con (4.23) y (4.24)

$$\begin{aligned} (b - d - f)g + (b + c - d)h &= 0; \\ (g + h)b + (b + c)f &= bg + bh + ch. \end{aligned}$$

Este sistema nos da como solución $g = b + c - d$ y $h = -b + d + f$ por lo que $g + h = c + f$. De forma análoga hacemos los otros dos casos:

Para (4.24) y (4.25) tenemos el siguiente sistema de ecuaciones.

$$\begin{aligned} (c - g - h)d + (b + c - g)f &= 0; \\ cd + (b + c)f &= bg + bh + ch; \end{aligned}$$

con solución $d = b + c - g$ y $f = -c + g + h$, a lo que nos da que $d + f = b + h$.

Por último (4.25) y (4.23) nos dan el siguiente sistema

$$\begin{aligned} (f - g - h)b + (d + f - h)c &= 0; \\ (g + h)b + hc &= 0; \end{aligned}$$

con solución $b = d + f - h$ y $c = -f + g + h$ y por lo tanto $b + c = d + g$.

Entonces la matriz de probabilidades de transición queda de la siguiente forma

$$P = \begin{pmatrix} 1 - b - c & b & c \\ d & 1 - d - f & f \\ g & h & 1 - g - h \end{pmatrix}$$

es doblemente estocástica pues $1 - b - c + d + g = 1 - b - c + b + c = 1$, $1 - d - f + b + h = 1 - d - f + d + f = 1$ y $1 - g - h + c + f = 1 - g - h + c + h = 1$. Hemos demostrado que si los invariantes del árbol generador de los sistemas de ciclos de primer retorno son iguales para cada vértice, entonces la matriz de probabilidades de transición P es doblemente estocástica.

Nos falta demostrar el regreso de la prueba, donde suponemos que la matriz es doblemente estocástica. Demostremos que todos los invariantes del árbol generador de todos los sistemas de ciclos primer retorno son iguales y con valor 3. Demostrémoslo para el primer vértice.

Tomemos la ecuación (4.16). Cabe mencionar que para sacar la cuenta hasta tal momento nunca utilizamos ninguna otra suposición más que P es estocástica. Supongamos que P es doblemente estocástica, entonces P cumple con las siguientes igualdades $b+c = d+g$, $b+h = d+f$ y $g+h = c+f$. Por lo que la ecuación (4.16) se simplifica de la siguiente forma

$$\begin{aligned}\tau(f_1(t)) &= 1 + \frac{fb + fc + bg + hb + hc + dc}{fg + hd + gd} = 1 + \frac{(f + g + h)b + (h + d + f)c}{fg + hd + gd} = \\ &= 1 + \frac{(2f + c)b + (2d + 2f - b)c}{fg + fd + cd} = 1 + 2\frac{bf + cd + cf}{fg + fd + cd} = 1 + 2\frac{fg + fd + cd}{fg + fd + cd} = 1 + 2 = 3.\end{aligned}$$

Para el segundo vértice tenemos que (4.17), de igual forma que (4.16), sólo es el valor del invariante del árbol generador en el sistemas de ciclos primer retorno del vértice 2 bajo la suposición que P es estocástica y ninguna otra.

$$\begin{aligned}\tau(f_2(t)) &= 1 + \frac{fb + fc + fg + dc + dg + dh}{gb + hb + hc} = 1 + \frac{f(b + c + g) + d(c + g + h)}{bg + bh + cf} = \\ &= 1 + \frac{f(d + 2g) + d(2g + 2h - f)}{bg + dh + gh} = 1 + 2\frac{bg + bh + ch}{bg + dh + gh} = 1 + 2\frac{bg + dh + gh}{bg + dh + gh} = 1 + 2 = 3.\end{aligned}$$

De igual forma para (4.18) tenemos que

$$\begin{aligned}\tau(f_3(t)) &= 1 + \frac{gb + gd + gf + hb + hc + hd}{dc + fb + fc} = 1 + \frac{g(b + d + f) + h(b + c + d)}{bf + cd + cf} = \\ &= 1 + \frac{g(2b + h) + h(2b + 2c - g)}{bc + bf + ch} = 1 + 2\frac{bg + bh + ch}{bc + bf + ch} = 1 + 2\frac{bc + bf + ch}{bc + bf + ch} = 1 + 2 = 3.\end{aligned}$$

Entonces para cada uno de los tres sistemas de ciclos primer retorno, tenemos que si la matriz es doblemente estocástica entonces el invariante del árbol generador en cada uno de los sistemas de ciclos primer retorno tienen valor 3. Por lo tanto son iguales. ✦

Para cada matriz estocástica P de tamaño $n \times n$, podemos determinar un vector $\tau_P \in \mathbb{R}^n$ el cual tiene en sus entradas el valor del invariante del árbol generador en el sistema de ciclos de primer retorno del respectivo vértice. Llamaremos a este vector el espectro del invariante de árbol generador en los ciclos de primer retorno. De esta forma se puede demostrar que cuando la entropía es máxima y la digráfica asociada a la matriz P es completa, entonces el espectro del invariante del árbol generador en los sistemas de ciclos de primer retorno tiene un sólo valor en todas sus entradas.

Capítulo 5

Conclusiones

Obtener una fórmula explícita del invariante del árbol en los sistemas de primer retorno no es una cuestión fácil como se ve, sin embargo con ayuda de la computadora se ve que la siguiente afirmación es cierta.

- *Sea $P \in \mathcal{M}_{n \times n}((0,1))$ una matriz de transición de probabilidades de una cadena de Markov. Entonces P es doblemente estocástica si y sólo si el invariante del árbol generador para los sistemas ciclos de primer retorno es igual para cada uno de los n vértices y con valor n .*

Sin embargo al finalizar esta Tesis no contamos con una prueba formal de tal afirmación. (Recientemente se ha reportado que los invariantes del árbol generador de los sistemas de ciclos de primer retorno son en realidad los conocidos “tiempos promedio de recurrencia”, en cuyo caso resulta verdadera la afirmación anterior. Más aún, recientemente Ma. Emilia Caballero observó que la serie de potencias que describe al sistema de ciclos de primer retorno en un vértice v es la función generadora de momentos de la variable aleatoria de los tiempos de primer retorno.) Pero el invariante del árbol generador aplicado a los sistemas de ciclos de primer retorno no ha sido explorado aún como un invariante de otro tipo de isomorfismos.

Podemos definir el espectro del invariante del árbol generador, sería simplemente un vector en \mathbb{R}^n , donde la entrada i del vector sería el valor del invariante del árbol generador aplicado al sistema de primer retorno del vértice i . Este vector suponemos que es un invariante bajo cierto tipo de isomorfismos.

Uno podría preguntarse el por qué definir el invariante de árbol generador en los sistemas de ciclos de primer retorno, si no cumple con ser un invariante bajo isomorfismos de palabra

mágica. La diferencia es que el invariante de árbol generador aplicado a la gráfica original no considera el comportamiento recurrente. Como los sistemas de ciclos de primer retorno tienen un comportamiento recurrente por construcción, entonces al aplicar el invariante del árbol generador a los sistemas de primer retorno, tenemos asociado a dicho invariante un comportamiento recurrente. De esta forma es posible considerar otros isomorfismos distintos a la conjugación para determinar si dos σ -espacios son isomorfos o no bajo dichos isomorfismos.

Bibliografía

- [1] R.B. Bapat, Raghavan. *Nonnegative matrices and applications*. (Encyclopedia of mathematics and its applications) Cambridge University Press, (1997)
- [2] G. Birkhoff. *Three observations on linear algebra*, Univ. Nac. Tucumán. Rev. Ser. A 5 (1946) 147-151.
- [3] M. Boyle, J. Buzzi, R. Gómez. Almost isomorphism of countable state Markov shifts. *Journal für die reine und angewandte Mathematik (Crelles's Journal)*. (2006), 592, 23-47.
- [4] M.E. Caballero, V.M Rivero, G. Uribe Bravo, C. Velarde *Cadenas de Markov: Un enfoque elemental*. (Aportaciones Matemáticas) **64** Sociedad Matemática Mexicana (2008).
- [5] P. Flajolet and ... *Analytic combinatorics*. Cambridge University Press (2009).
- [6] N. Friedman and D.S. Ornstein. On the isomorphism of weak Bernoulli transformations. *Advances in Math.* **5** (1970) 365-394.
- [7] F.R. Gantmacher. *Matrix Theory, Volume II*. Chelsea Publishing Univ. Press, Princeton (1960).
- [8] G.R. Grimmett y D.R. Stirzaker. *Probability and Random Processes*. Oxford University Press. 3a edición. (2001).
- [9] R. Gómez. Positive K-theory for finitary isomorphisms of Markov chains. *Erog. Th. and Dyn. Sys.* (2003), 23, 1485-1504.
- [10] B.M. Gurevich and S.V. Savchenko. Thermodynamic formalism for countable symbolic Markov chains. *Russian Math. Surveys* **53:2** (1998) 3-106.
- [11] T.W. Hungerford. *Algebra*. Springer (1974).
- [12] D. Lind and B. Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press (1995).

-
- [13] D. Lind and S. Tuncel. A spanning tree invariant for Markov shifts Codes, Systems, and Graphical Models (Minneapolis, MN, 1999), volume 123 of IMA Vol. Math. Appl.
- [14] B. Marcus and S. Tuncel. The weight-per-symbol polytope and scaffolds of invariants associated with Markov chains. *Ergod. Th. and Dynam. Sys.* **11** (1991) 129-180.
- [15] D.S. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Math.* **4** (1970) 337-352.
- [16] B. Parry. A finitary classification of topological Markov chains and sofic systems. *Bull. London Math. Soc.* **9** (1977) 86-92.
- [17] H.L. Royden. Real Analysis. Macmillan (1988).
- [18] C. Shannon. A mathematical theory of communication. Bell Sys. Tech. J. 27 (1948), 379-423, 623-656.
- [19] J.M. Salazar Homomorfismos en σ -espacios de Cantor Weyl y Besicovitch. Tesis Licenciatura UNAM. (2007).
- [20] P. Walters, *An Introduction to Ergodic Theory*, Springer graduated Text in Maths **79**, New York, 1982
- [21] R.F. Williams Clasificación de subshifts of finite tipe, *Annals of Maths* **98** (1973) 120-153
erratum *Annals of maths* **99** (1974)341-361